



PCT / IB 98 / 0 0 6 2 5

2 4. 04. 98

09 / 40 3 7 2 4

SCHWEIZERISCHE EIDGENOSSENSCHAFT  
CONFÉDÉRATION SUISSE  
CONFEDERAZIONE SVIZZERA

REC'D	29 APR 1998
WIPO	PCT

### Bescheinigung

Die beiliegenden Akten stimmen mit den ursprünglichen technischen Unterlagen des auf der nächsten Seite bezeichneten Patentgesuches für die Schweiz und Liechtenstein überein. Die Schweiz und das Fürstentum Liechtenstein bilden ein einheitliches Schutzgebiet. Der Schutz kann deshalb nur für beide Länder gemeinsam beantragt werden.

### Attestation

Les documents ci-joints sont conformes aux pièces techniques originales de la demande de brevet pour la Suisse et le Liechtenstein spécifiée à la page suivante. La Suisse et la Principauté de Liechtenstein constituent un territoire unitaire de protection. La protection ne peut donc être revendiquée que pour l'ensemble des deux Etats.

PRIORITY DOCUMENT

### Attestazione

Gli uniti documenti sono conformi agli atti tecnici originali della domanda di brevetto per la Svizzera e il Liechtenstein specificata nella pagina seguente. La Svizzera e il Principato di Liechtenstein formano un unico territorio di protezione. La protezione può dunque essere rivendicata solamente per l'insieme dei due Stati.

Bern, - 3. April 1998

Eidgenössisches Institut für Geistiges Eigentum  
Institut Fédéral de la Propriété Intellectuelle  
Istituto Federale della Proprietà Intellettuale

Patentgesuche  
Demandes de brevet  
Domande di brevetto

*U. Kohler*

1980



Patentgesuch Nr. 1997 0966/97

HINTERLEGUNGSBESCHEINIGUNG (Art. 46 Abs. 5 PatV)

Das Eidgenössische Institut für Geistiges Eigentum bescheinigt den Eingang des unten näher bezeichneten schweizerischen Patentgesuches.

Titel:  
Neurotrypsin.

Patentbewerber:  
Prof. Dr. Peter Sonderegger Biochemisches Institut Universität Zürich  
Winterthurerstrasse 190  
8057 Zürich

Vertreter:  
Patentanwaltsbüro Zink  
Birchlistrasse 11  
8173 Riedt-Neerach

Anmeldedatum: 26.04.1997

Voraussichtliche Klassen: A61K, C07K, C12N, C12P



### Neurotrypsin

Die vorliegende Erfindung betrifft Neurotrypsine und ein Medikament, welches diese Substanzen enthält oder auf diese Substanzen einwirkt

5

Neurotrypsin ist eine neu entdeckte Serinprotease, welche vor allem im Gehirn und in der Lunge exprimiert wird; die Expression im Gehirn findet fast ausschliesslich in Nervenzellen statt.

10

Neurotrypsin hat eine bisher nicht gefundene Domänenzusammensetzung: Neben der Protease-Domäne findet man 3 oder 4 SRCR (Scavenger Receptor Cysteine-Rich)-Domänen und eine Kringle-Domäne. Es ist hervorzuheben, dass die Kombination von Kringle- und SRCR-Domänen bisher noch nie in Proteinen gefunden wurde. Am Aminotermius des Neurotrypsin-Proteins befindet sich ein Segment von

15

über 60 Aminosäuren, welches einen ausserordentlich hohen Anteil von Prolin und basischen Aminosäuren (Arginin und Histidin) aufweist.

Die Erfindung ist durch die Merkmale in den unabhängigen Ansprüchen gekennzeichnet. Bevorzugte Ausführungsformen sind in den abhängigen Ansprüchen

20

definiert.

Die neu gefundenen Neurotrypsine

- Neurotrypsin des Menschen (Verbindung der Formel I),

25

- Neurotrypsin der Maus (Verbindung der Formel II)

unterscheiden sich strukturell sehr stark von den bisher bekannten Serinproteasen.

30

Die bezüglich der Protease-Domäne strukturell am nächsten mit den neuen Verbindungen verwandte Serinprotease, nämlich Plasmin (des Menschen), weist eine Aminosäuresequenz-Identität von lediglich 44% auf.

Das Prolin-reiche, basische Segment am Aminotermminus weist eine gewisse  
Aehnlichkeit auf zu basischen Segmenten der Netrine und der  
Semaphorine/Collapsine. Aufgrund dieses Segmentes ist es wahrscheinlich, dass  
5 Neurotrypsin mittels Heparin-Affinitätschromatographie angereichert werden kann.

Die Neurotrypsin des Menschen (Verbindung der Formel I), und der Maus  
(Verbindung der Formel II) weisen unter sich eine sehr hohe strukturelle Ähnlichkeit  
auf.

10

Die Identität der Aminosäuresequenzen der nativen Proteine der  
Verbindungen der Formeln I oder II beträgt 81%.

Das Neurotrypsin des Menschen (Verbindung der Formel I) weist eine  
15 codierende Sequenz von 2625 Nucleotiden auf. Das codierte Peptid der Verbindung  
der Formel I ist 875 Aminosäuren lang und enthält ein Signalpeptid von 20 Amino-  
säuren. Das Neurotrypsin der Maus (Verbindung der Formel II) weist eine codierende  
Sequenz von 2283 Nucleotiden auf. Das codierte Protein der Verbindung der  
Formel II ist 761 Aminosäuren lang und enthält ein Signalpeptid von 21 Aminosäuren.  
20 Der Grund für die grössere Länge des Neurotrypsins des Menschen liegt darin, dass  
dieses 4-SRCR-Domänen aufweist, während das Neurotrypsin der Maus nur  
3 SRCR-Domänen hat.

Die bei beiden Verbindungen (Verbindung der Formel I und Verbindung der  
25 Formel II) vorhandenen Domänen weisen einen hohen Grad von Sequenzähnlichkeit  
auf. Die einander entsprechenden SRCR-Domänen der Verbindungen der Formeln I  
und II weisen eine Aminosäuresequenzidentität von 81% bis 91% auf. Die  
entsprechenden Kringle-Domänen haben eine Aminosäuresequenzidentität von 75%.  
Ein hoher Grad von Aehnlichkeit besteht vor allem auch in der enzymatisch aktiven  
30 (d.h. proteolytischen) Domäne (90% Aminosäuresequenzidentität).

Die Proteasedomänen der Neurotrypsine des Menschen (Verbindung der Formel I) und der Maus (Verbindung der Formel II) sind im Folgenden gegeneinander aufgereiht, um den hohen Grad von Sequenzidentität zu illustrieren.

```
CGLRLLHRRQKRIIGGKNSLRGGWPWQVSLRLKSSHGDGRLLCGATLLSS      50
|||||:|||||.||||:|.|||||
CGLRLLHRRQKRIIGGNNSLRGAWPWQASLRLRSAHGDGRLLCGATLLSS
```

```
CWVLTAAHCFKRYGNSTRSYAVRVGDYHTLVPEEFEEEEIGVQQIVIHREY      100
|||||:|||||.||||:|.|||||
CWVLTAAHCFKRYGNNSRSYAVRVGDYHTLVPEEFEQEIGVQQIVIHREY
```

```
RPDRSDYDIALVRLQGPEEQCARFSSHVLPACLPLWRERPQKTASNCYIT      150
|||||:|||||.||||:|.|||||
RPDRSDYDIALVRLQGPGEQCARLSTHVLPACLPLWRERPQKTASNCHIT
```

```
GWGDTGRAYSRTLQQAAILLPKRFCCEERYKGRFTGRMLCAGNLHEHKRV      200
|||||:|||||.||||:|.|||||
GWGDTGRAYSRTLQQAAVPLLPKRFCCKERYKGLFTGRMLCAGNLQEDNRV
```

```
DSCQGDSSGGPLMCERPGESWVVYGVTSWGYGCGVKDSPGVYTKVSAFVPW      250
|||||:|||||.||||:|.|||||
DSCQGDSSGGPLMCEKPDSESWVVYGVTSWGYGCGVKDTPGVYTRVPAFVPW
```

```
IKSVTKL      258
|||||.|
IKSVTSL
```

- 5 Von den in den Vergleich einbezogenen 258 Aminosäuresequenzpositionen sind 233 in beiden Verbindungen (Verbindung der Formel I und Verbindung der Formel II) mit der gleichen Aminosäure besetzt (markiert mit senkrechten Strichen).

Verglichen mit den bekannten Serinproteasen ist für die erfindungsgemässen Neurotrypsine einzigartig, dass sie gemäss bisherigen Erkenntnissen in ausgeprägtem Masse von Nervenzellen exprimiert werden. Ein anderes Organ mit starker Expression von Neurotrypsin ist die Lunge (siehe  
5 Gschwend et al., Mol. Cell. Neurosci., in press, 1997).

Die den Strukturen der Verbindungen der Formeln I oder II am stärksten gleichenden Proteine sind Serin-Proteasen, wie Gewebe-Plasminogenaktivator (tPA), Urokinase-Plasminogenaktivator (uPA), Plasmin, Trypsin, Apolipoprotein (a),  
10 Coagulation-Factor XI, Neuropsin, und Acrosin.

Im erwachsenen Gehirn werden die erfindungsgemässen Verbindungen vorwiegend in der Grosshirnrinde, dem Hippocampus und der Amygdala exprimiert.

15 Im erwachsenen Hirnstamm und Rückenmark werden die erfindungsgemässen Verbindungen vorwiegend in den motorischen Nervenzellen exprimiert. Eine etwas schwächere Expression ist in den Nervenzellen der oberflächlichen Schichten des Hinterhorns des Rückenmarks zu finden.

20 Im erwachsenen peripheren Nervensystem werden die erfindungsgemässen Verbindungen in einer Subpopulation der Spinalganglienneurone exprimiert.

Das Gen-Expressionsmuster der erfindungsgemässen Verbindungen im Gehirn ist äusserst interessant, weil diese Moleküle im adulten Nervensystem vor  
25 allem in Nervenzellen derjenigen Regionen exprimiert werden, denen eine wichtige Rolle bei Lern- und Gedächtnisfunktionen zugeschrieben wird.

Das Gen-Expressionsmuster der erfindungsgemässen Verbindungen in der Grosshirnrinde (vor allem Schicht V und VI) ist äusserst interessant, weil eine  
30 Reduktion der zellulären Differenzierung in der Grosshirnrinde in Assoziation mit Schizophrenie gefunden wurde.



Eine andere hervorzuhebende Eigenschaft der erfindungsgemässen Verbindungen besteht darin, dass sie von den Nervenzellen sezerniert werden.

5 Diese Tatsache - zusammen mit der Funktion als Protease und dem Expressionsmuster im sich entwickelnden und adulten Gehirn - lassen eine Rolle der erfindungsgemässen Verbindungen bei der Regulation der extrazellulären Proteolyse in Gehirnarealen vermuten, welche an der Verarbeitung und Speicherung von erlernten Verhaltensweisen, erlernten Gefühlen oder von Gedächtnisinhalten beteiligt sind.

10 Zusammen mit kürzlich gefundener Evidenz für eine Rolle von extrazellulären Proteasen bei der neuronalen Plastizität, lässt das Expressionsmuster vermuten, dass die proteolytische Wirkung von Neurotrypsin eine Rolle innehat bei strukturellen Reorganisationen im Rahmen von Lern- und Gedächtnis-Operationen, zum Beispiel Operationen, welche der Verarbeitung und Speicherung von erlernten Verhaltensweisen, erlernten Gefühlen oder von Gedächtnisinhalten beteiligt sind.

20 Das Gen-Expressionsmuster der erfindungsgemässen Verbindungen im Rückenmark und in den Spinalganglien ist interessant, weil diese Moleküle im adulten Nervensystem in Nervenzellen derjenigen Zellgruppierungen exprimiert werden, denen eine Rolle bei der Verarbeitung von Schmerz, sowie bei der Entstehung pathologischer Schmerzzustände zugeschrieben wird.

25 Die erfindungsgemässen Verbindungen wurden im Rahmen einer Studie gefunden, welche zum Ziel hatte, Trypsin-ähnliche Serinproteasen im Nervensystem aufzuspüren.

Als erste wurde die Verbindung der Formel II entdeckt und charakterisiert (siehe Gschwend et al., Mol. Cell. Neurosci., in press, 1997).

30

Durch ein "Alignment" der Proteasedomänen von 7 bekannten Serinproteasen (tissue-type Plasminogenaktivator, urokinase-type Plasminogenaktivator, Thrombin, Plasmin, Trypsin, Chymotrypsin und pankreatische Elastase) in den

Regionen des Histidins und des Serins der katalytischen Triade der aktiven Stelle wurden die Sequenzen von sogenannten "Primer-Oligonucleotiden" für die Polymerasen-Kettenreaktion ermittelt.

5 Die Primer-Oligonucleotide wurden in einer Polymerasen-Kettenreaktion (PCR) zusammen mit ss-cDNA aus Total-RNA aus Gehirnen von 10 Tage alten Mäusen eingesetzt und führten zur Amplifizierung eines cDNA-Fragments mit einer Länge von ungefähr 500 Basenpaaren.

10 Dieses cDNA-Fragment wurde erfolgreich zur Isolierung von weiteren cDNA-Fragmenten aus im Handel erhältlichen cDNA-Bibliotheken eingesetzt. Zusammen erstreckten sich die isolierten cDNA-Fragmente über die volle Länge des codierten Teils der Verbindung der Formel II.

15 Durch herkömmliche DNA-Sequenzierung wurde die vollständige Nucleotidsequenz und die davon abgeleitete Aminosäuresequenz der Verbindung der Formel II erhalten.

20 Die Verbindung der Formeln I wurde aufgrund ihrer ausgeprägten Ähnlichkeit mit der Verbindung der Formel II mittels herkömmlicher PCR kloniert.

Die eingesetzten Primer-Oligonucleotide wurden gemäss der bekannten Sequenz der Verbindung der Formel II synthetisiert.

25 Die Klonierung der Verbindung der Formel I wurde mittels zwei im Handel erhältlichen cDNA-Bibliotheken aus fötalem menschlichen Gehirn durchgeführt.

30 Diese Art der Klonierung kann auch zur Isolierung der homologen Verbindung anderer Spezies, wie Ratte, Kaninchen, Meerschweinchen, Rind, Schaf, Schwein, Primaten, Vögel, Zebrafisch (*Brachydanio rerio*), *Drosophila melanogaster*, *Caenorhabditis elegans* etc. verwendet werden.

Die codierenden Nucleotidsequenzen können eingesetzt werden zur Erzeugung von Proteinen mit den codierten Aminosäuresequenzen der Verbindungen der Formeln I oder II. Ein in unserem Labor praktiziertes Verfahren erlaubt die Produktion von rekombinanten Proteinen in Myelomazellen als Fusionsprotein mit einer Immunoglobulin-Domäne (konstante Domäne der Leichten-Kette-Kappa). Das Konstruktionsprinzip ist im Detail beschrieben durch Rader et al. (Rader et al., Eur. J. Biochem. 215, Seiten 133-141, 1993). Das so von den Myelomazellen synthetisierte Fusionsprotein wurde durch Immunoaffinitätschromatographie mittels eines monoklonalen Antikörpers gegen die Ig-Domäne der Leichten-Kette-Kappa isoliert. Mit der gleichen Expressionsmethode kann auch das native Protein einer Verbindung, ausgehend von der codierenden Sequenz, produziert werden.

Die codierenden Sequenzen der Verbindungen der Formeln I oder II können als Ausgangsverbindungen dienen für das Aufspüren und die Isolierung von Allelen der Verbindungen der Formeln I oder II. Sowohl die Polymerasen-Kettenreaktion, als auch die Nucleinsäure-Hybridisierung können zu diesem Zweck eingesetzt werden.

Die codierenden Sequenzen der Verbindungen der Formeln I oder II können als Ausgangsverbindungen dienen für sogenannte "site-directed mutagenesis", um Nucleotidsequenzen zu generieren, welche die durch die Verbindungen der Formeln I oder II codierten Proteine, oder Teile davon, codieren, aber als Resultat unterschiedlicher Verwendung alternativer Codons, degeneriert sind bezüglich der als Verbindungen der Formel I oder II definierten Nucleotidsequenzen.

Die codierenden Sequenzen der Verbindungen der Formeln I oder II können als Ausgangsverbindungen dienen für die Herstellung von Sequenzvarianten durch sogenannte "site-directed mutagenesis".

Die nachfolgenden Beispiele illustrieren die vorliegende Erfindung.

Beispiel 1:

# cDNA-Klonierung der Verbindung der Formel II (Neurotrypsin der Maus)

- 5            Totale RNA wurde aus dem Gehirn von 10 Tage alten Mäusen (ICR-ZUR) gemäss der Methode von Chomczynski und Sacchi (1987) isoliert. Herstellung von einzelsträngiger cDNA erfolgte unter Benützung von Oligo(dT)-"Primer" und einer RNA-abhängigen DNA-Polymerase (SuperScript RNase H<sup>-</sup>Reverse Transcriptase; Gibco-BRL, Gaithersburg, MD) gemäss den Instruktionen des Herstellers. Für die
- 10    Durchführung der Polymerasen-Kettenreaktion wurden ein "Primer" in Leserichtung gemäss einem Aminosäuresequenzabschnitt aus der Region des konservierten Histidins der katalytischen Triade und ein "Primer" in entgegengesetzter Richtung gemäss einem Aminosäuresequenzabschnitt aus der Region des konservierten Serins der katalytischen Triade der Serinproteasen synthetisiert. Die für die
- 15    Bestimmung der Oligonucleotid-Primer eingesetzten Aminosäuresequenzen und deren Ermittlung durch den Vergleich der Sequenzen von 7 bekannten Serinproteasen ist im folgenden dargestellt.

Protease domain		I →		← II	
		N - [ H ] [ D ] - C		[ S ] - C	
tPA (m)	..SSC	W V L S A A H C	FLE.....HDA	C Q G D S G G	PLV..
uPA (m)	..SPC	W V A S A A H C	FIQ.....TDS	C K G D S G G	PLI..
thrombin (m)	..SDR	W V L T A A H C	ILY.....GDA	C E G D S G G	PFV..
plasmin (m)	..APE	W V L T A A H C	LKS.....VDS	C Q G D S G G	PLV..
trypsin (m)	..NDQ	W V V S A A H C	YKY.....KDS	C Q G D S G G	PVV..
chymTryp b (r)	..SED	W V V T A A H C	GVK.....VSS	C M G D S G G	PLV..
pancElas II (m)	..ANN	W V L T A A H C	LSN.....TSS	C N G D S G G	PLN..

Primer      (I) 5'-TGG GTI SYI WSI GCI GCI CAT TG-3'    (II) 3'-ACR BTY CCI CTR WSI CCI CC-5'

- Die Proteasen-Domänen von 7 bekannten Serinproteasen (tissue-type-Plasminogenaktivator, Urokinase-type Plasminogenaktivator, Thrombin, Plasmin,
- 20    Trypsin, Chymotrypsin und pankreatische Elastase) wurden im Bereich des konservierten Histidins und Serins der katalytischen Triade der aktiven Stelle gegeneinander aufgereiht. Die in diesen Regionen konservierten Aminosäuren wurden als Ausgangssequenz für die Bestimmung von degenerierten Primern

benutzt. Die Primersequenzen sind nach der Empfehlung der IUB-Nomenklatur (Nomenclature Committee, 1985) angegeben.

Die in der PCR eingesetzten Primer trugen zur Erleichterung einer späteren Klonierung zusätzlich die Restriktionsstellen *EcoRI* und *BamHI* am 5'-Ende.

5

Folgende Primer wurden eingesetzt:

In Leserichtung (sense primers):

5'-GGGGAATTCTGGGT(C/G)(T/C)I(T/A)(G/C)IGCIGCICA(T/C)TG-3'

10

In Gegenrichtung (antisense primers):

5'-GGGGGATCCCCICCI(G/C)(A/T)(A/G)TCICC(C/T)T(G/C/T)(G/A)CA-3'.

Die Polymerasen-Kettenreaktion wurde unter Standard-Bedingungen mittels der DNA-Polymerase AmpliTaq (Perkin Elmer) gemäss den Empfehlungen des  
15 Produzenten durchgeführt. Das folgende PCR-Profil wurde eingesetzt: 93°C für 3 Minuten, gefolgt von 35 Zyklen von 93°C für 1 Minute, 48°C für 2 Minuten und 72°C für 2 Minuten. Im Anschluss an den letzten Zyklus wurde die Inkubation bei 72°C während weiteren 10 Minuten fortgesetzt.

20

Die amplifizierte Fragmente hatten eine Länge von ungefähr 500 Basenpaaren. Sie wurden mit *EcoRI* und *BamHI* geschnitten und in einen Bluescript-Vektor eingefügt (Bluescript SK(-), Stratagene). Die resultierenden Klone wurden durch DNA-Sequenzbestimmung mittels der Dideoxy-Kettenterminations-Methode (Sanger et al., Proc. Natl. Acad. Sci. USA 77, 2163-2167, 1977) auf einem  
25 automatisierten DNA-Sequenziergerät (LI-COR, Modell 4000L; Lincoln, NE) unter Benützung eines kommerziellen Sequenzierkits (SequiTherm long-read cycle sequencing kit-LC; Epicentre Technologies, Madison, WI) analysiert. Die Analyse führte zu einer 474 Basenpaare umfassenden Sequenz aus dem katalytischen Bereich der Serinproteasedomäne der Verbindung der Formel II.

30

Das 474 Basenpaar lange PCR-Fragment wurde zum Absuchen einer Oligo(dT)-"primed" Uni-ZAP-XR-cDNA-Bibliothek aus Gehirn von 20 Tage alten Mäusen (Stratagene; Cat. No. 937 319) eingesetzt. Es wurden  $3 \times 10^6$  Lambda-

Plaques mittels des radioaktiv markierten PCR-Fragments unter hochstringenten Bedingungen (Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, 1989) abgesucht und 24 positive Klone gefunden und isoliert.

5

Aus den positiven Lambda-Uni-ZAP-XR-Phagemid-Klonen wurde das entsprechende Bluescript-Plasmid nach der Standardmethode gemäss den Empfehlungen des Herstellers (Stratagene) durch *in vivo*-Exzision herausgeschnitten. Um die Länge der eingefügten Fragmente zu bestimmen, wurden die entsprechenden Bluescript-Plasmid-Klone mit *SacI* und *KpnI* verdaut. Die Klone, welche die längsten Fragmente enthielten, wurden mittels DNA-Sequenzierung (wie oben beschrieben) und anschliessender Auswertung mittels der GCG-Software (Version 8.1, Unix; Silicon Graphics, Inc.) analysiert.

10

Da keiner der Klone die codierende Sequenz in voller Länge enthielt, wurde eine zweite cDNA-Bibliothek abgesucht. Die hierfür eingesetzte Bibliothek war eine Oligo(dT)- und "Random-Primed" cDNA Bibliothek in Lambda-Phagen (Lambda gt10), welche auf mRNA aus 15 Tage alten Maus-Embryonen basierte (oligo(dT)- and random-primed Lambda gt10 cDNA library; Clontech, Palo Alto, CA; Kat. No. ML 3002a). Als Sonde wurde ein radioaktiv markiertes DNA-Fragment (*AvaI*/*AatII*) vom 5'-Ende des längsten Klones der ersten Suche eingesetzt, und es wurden ungefähr  $2 \times 10^6$  Plaques abgesucht. Dabei wurden 14 positive Klone gefunden. Die cDNA-Fragmente wurden mittels *EcoRI* herausgeschnitten und in den Bluescript-Vektor (KS(+); Stratagene) kloniert. Die Sequenzanalyse erfolgte wie oben beschrieben.

20

25

Man erhielt so die Nucleotidsequenz über die volle Länge der cDNA von 2361 resp. 2376 Basenpaaren der Verbindung der Formel II. Mit dem beschriebenen Verfahren der PCR-Klonierung können auch Varianten-Formen der Verbindungen der Formeln I und II gefunden und isoliert werden, beispielsweise deren Allele, oder deren Splice-Varianten. Das beschriebene Verfahren des Absuchens einer cDNA-Bibliothek ermöglicht auch das Auffinden und die Isolierung von Verbindungen, welche unter stringenten Verbindungen an die codierenden Sequenzen der Formeln I und II hybridisieren.

30

Beispiel 2:

5 Klonierung der cDNA der Verbindung der Formel I (Neurotrypsin des Menschen)

Die Klonierung der cDNA der Verbindung der Formel I wurde auf der Grundlage der Nucleotid-Sequenz der Verbindung der Formel II durchgeführt. Als erster Schritt wurde ein Fragment der Verbindung der Formel I mittels  
10 Polymerasenkettenreaktion (PCR) amplifiziert. Als Matrize diente die DNA aus einer cDNA-Bibliothek vom Gehirn eines menschlichen Foetus (17. - 18. Schwangerschaftswoche), welche auf dem Markt (Oligo(dT)- and random-primed, human fetal brain cDNA library in the Lambda ZAP II vector, Cat. No. 936206, Stratagene) erhältlich ist. Die synthetischen PCR-Primer enthielten, zur Erleichterung  
15 der nachfolgenden Klonierung, die Restriktionsstellen Hind III und Xho I am 5'-Ende.

In Leserichtung (sense primers):

5'-GGGAAGCTTGGICA(A/G)TGGGGIACI(A/G)TITG(C/T)GA(C/T)-3'

20 In Gegenrichtung (antisense primer):

5'-GGGCTCGAGCCCCAICCTGTTATGTAAIAGTTG-3'

Die PCR wurde unter Standard-Bedingungen mittels der DNA-Polymerase  
25 Amplitaq (Perkin Elmer) gemäss den Empfehlungen des Produzenten durchgeführt. Das entstandene Fragment von 1116 Basenpaaren wurde in einen Bluescript-Vektor eingefügt (Bluescript SK(-), Stratagene). Ein 600 Basenpaare langes Hind III/Stu I-Fragment, entsprechend der 5'-Hälfte des 1116 Basenpaare langen PCR-Fragments, wurde zum Absuchen einer Lambda-cDNA-Bibliothek aus menschlichem foetalem  
30 Gehirn (Human Fetal Brain 5'-STRETCH PLUS cDNA library; Lambda gt10; Cat. No. HL3003a; Clontech) eingesetzt. Es wurden  $2 \times 10^6$  Lambda-Plaques mittels des radioaktiv markierten PCR-Fragments unter hochstringenten Bedingungen

(Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, 1989) abgesucht und 23 positive Klone gefunden und isoliert.

Aus den positiven Lambda-gt10-Klonen wurden die entsprechenden cDNA-  
 5 Fragmente mit *EcoRI* herausgeschnitten und in einen Bluescript-Vektor eingefügt  
 (Bluescript KS(+), Stratagene). Die Sequenzierung erfolgte mittels der Dideoxy-  
 Kettenterminations-Methode (Sanger et al., Proc. Natl. Acad. Sci. USA 77, Seiten  
 2163-2167, 1977), unter Verwendung eines kommerziellen Sequenzierkits (Sequi  
 Therm long-read cycle sequencing kit-LC; Epicentre Technologies, Madison, WI) und  
 10 Bluescript-spezifischen Primern.

In einer alternativen Sequenzier-Strategie wurden die cDNA-Fragmente der  
 positiven Lambda-gt10-Klone, unter Verwendung Lambda-spezifischer Primer, mittels  
 PCR amplifiziert. Die Sequenzierung wurde wie oben beschrieben durchgeführt.  
 15

Die computerisierte Analyse der Sequenzen wurde mittels des  
 Programmpakets GCG (Version 8.1, Unix; Silicon Graphics Inc.) durchgeführt.

Man erhielt so die Nucleotid-Sequenz über die volle Länge der cDNA von  
 20 3350 Basenpaaren. Mit dem beschriebenen Verfahren der PCR-Klonierung können  
 auch Varianten-Formen der Verbindungen der Formeln I oder II gefunden und isoliert  
 werden, beispielsweise deren Allele, oder deren Splice-Varianten. Das beschriebene  
 Verfahren des Absuchens einer cDNA-Bibliothek ermöglicht auch das Auffinden und  
 die Isolierung von Verbindungen, welche unter stringenten Bedingungen an die  
 25 codierenden Sequenzen der Formeln I oder II hybridisieren.



Beispiel 3:

Nachweis der codierten Sequenzen der Verbindungen I oder II mittels Antikörpern

5

Das mehr als 60 Aminosäuren lange Prolin-reiche, basische Segment am Aminotermius der codierten Sequenz der Verbindungen der Formeln I oder II eignet sich gut für die Herstellung von Antikörpern mittels der Synthese von Peptiden und deren Einsatz zur Immunisierung. Wir haben aus dem Prolin-reichen, basischen  
10 Segment am Aminotermius der codierten Sequenz der Verbindung der Formel II zwei Peptidsequenzen mit einer Länge von 19, respektive 13, Aminosäuren zur Erzeugung von Antikörpern ausgewählt. Die Peptide hatten folgen Sequenzen:

Peptid 1:  $\text{H}_2\text{N-SRS PLH RPH PSP PRS QX-CONH}_2$

Peptid 2:  $\text{H}_2\text{N-LPS SRR PPR TPR F-COOH}$

15

Die beiden Peptide wurden chemisch synthetisiert, an eine makromolekulare Trägersubstanz (Keyhole Limpet Hemacyanin) gekoppelt, und zur Immunisierung von 2 Kaninchen injiziert. Die erzeugten Antiseren wiesen einen hohen Antikörper-Titer auf und konnten erfolgreich sowohl zur Identifizierung von nativem Neurotrypsin aus  
20 Gehirnextrakt der Maus als auch zur Identifizierung von recombinantem Neurotrypsin eingesetzt werden. Das angewandte Verfahren zur Erzeugung von Antikörpern kann auch zur Erzeugung von Antikörpern gegen die codierte Sequenz der Verbindung der Formel I angewendet werden.

25

Die erzeugten Antikörper gegen die Teilsequenzen der codierten Sequenzen der Formeln I oder II können zur Aufspürung und zur Isolierung von Varianten-Formen der Verbindungen der Formeln I oder II, wie beispielsweise Allele oder Splice-Varianten, eingesetzt werden. Auch gentechnisch erzeugte Varianten der Verbindungen der Formeln I oder II können mit solchen Antikörpern aufgespürt und  
30 isoliert werden.

Beispiel 4:

**Reinigung der codierten Sequenzen der Verbindungen der Formeln I oder II**

5           Zur Reinigung der codierten Sequenzen der Verbindungen der Formeln I oder II können, neben konventionellen chromatographischen Methoden, wie beispielsweise Ionenaustauscher-Chromatographie, zwei affinitätschromatographische Methoden eingesetzt werden. Eine der affinitätschromatographischen  
10       Reinigungsprozeduren basiert auf der Verfügbarkeit von Antikörpern. Durch Kopplung der Antikörper an eine chromatographische Matrix kann ein Reinigungsverfahren erzeugt werden, das in einem Schritt einen sehr hohen Grad an Reinheit der entsprechenden Verbindung verspricht.

15           Ein für die Reinigung der codierten Sequenzen der Verbindungen der Formeln I und II ebenfalls wichtiges Merkmal ist das Prolin-reiche, basische Segment am Aminotermminus. Es ist zu erwarten, dass, aufgrund der hohen Dichte an positiven Ladungen, dieses Segment die Bindung der codierten Sequenzen der Verbindungen der Formeln I oder II an Heparin und Heparin-ähnliche Affinitätsmatrices vermittelt. Dieses Prinzip ermöglicht auch die Isolierung, oder zumindest die Anreicherung von  
20       Varianten-Formen der codierten Sequenzen der Formeln I oder II, beispielsweise deren Allele oder Splice-Varianten. Gleichermassen kann die Heparin-Affinitätschromatographie auch zur Isolierung, oder zumindest zur Anreicherung, von speziesshomologen Proteinen der Verbindungen der Formeln I oder II eingesetzt werden.

25

          Nachfolgend werden Angaben zu den Verbindungen der Formeln I oder II gemacht.

30

- (1) ANGABEN ZUR VERBINDUNG DER FORMEL I (Neurotrypsin des Menschen)
- (i) SEQUENZKENNZEICHEN:
- 5
- (A) LÄNGE: 3350 Basenpaare  
(B) ART: Nucleinsäure  
(C) STRANGFORM: Einzelstrang  
(D) TOPOLOGIE: linear
- 10
- (ii) ART DES MOLEKÜLS: c-DNA zu m-RNA
- (vi) URSPRÜNGLICHE HERKUNFT:
- 15
- (A) ORGANISMUS: Homo sapiens  
(D) ENTWICKLUNGSSTADIUM: Foetalstadium  
(F) GEWEBETYP: Gehirn
- (vii) UNMITTELBARE HERKUNFT:
- 20
- (A) BIBLIOTHEK: human fetal brain 5'-stretch plus cDNA library in the lambda gt10 vector; catalog No. HL 3003a; clontech, Palo Alto, CA, USA.
- 25
- (B) CLONE: cDNA-Klone No.:  
3-1, 3-2, 3-6, 3-7, 3-8, 3-10, 3-11, 3-12

(ix) MERKMAL:

(A) NAME/SCHLÜSSEL: Signalpeptid

(B) LAGE: 44 .. 103

5

(ix) MERKMAL:

(A) NAME/SCHLÜSSEL: reifes Peptid

10 (B) LAGE: 104 .. 2668

(ix) MERKMAL:

15 (A) NAME/SCHLÜSSEL: codierende Sequenz

(B) LAGE: 44 .. 2668

(ix) MERKMAL:

20

(A) NAME/SCHLÜSSEL: Protein-reiches basisches Segment

(B) LAGE: 104 .. 319

25 (ix) MERKMAL:

(A) NAME/SCHLÜSSEL: Kringle-Domäne

(B) LAGE: 320 .. 538

30

(ix) MERKMAL:

(A) NAME/SCHLÜSSEL: SRCR-Domäne 1

(B) LAGE: 551 .. 856

5

(ix) MERKMAL:

(A) NAME/SCHLÜSSEL: SRCR-Domäne 2

10 (B) LAGE: 881 .. 1186

(ix) MERKMAL:

15 (A) NAME/SCHLÜSSEL: SRCR-Domäne 3

(B) LAGE: 1202 .. 1504

(ix) MERKMAL:

20

(A) NAME/SCHLÜSSEL: SRCR-Domäne 4

(B) LAGE: 1541 .. 1846

25 (ix) MERKMAL:

(A) NAME/SCHLÜSSEL: proteolytische Domäne

(B) LAGE: 1898 .. 2668

30

(ix) MERKMAL:

(A) NAME/SCHLÜSSEL: Histidin der katalytischen Triade

(B) LAGE: 2069 - 2071

5

(ix) MERKMAL:

(A) NAME/SCHLÜSSEL: Asparaginsäure der katalytischen Triade

10 (B) LAGE: 2219 - 2221

(ix) MERKMAL:

15 (A) NAME/SCHLÜSSEL: Serin der katalytischen Triade

(B) LAGE: 2516 .. 2518

(ix) MERKMAL:

20

(A) NAME/SCHLÜSSEL: PolyA-Signal

(B) LAGE: 2873 .. 2878

25 (ix) MERKMAL

(A) NAME/SCHLÜSSEL: PolyA-Signal

(B) LAGE: 3034 .. 3039

30

(ix) MERKMAL:

- (A) NAME/SCHLÜSSEL: PolyA-Signal
- (B) LAGE: 3215 .. 3220

5 (ix) MERKMAL

- (A) NAME/SCHLÜSSEL: 3'UTR
- (B) LAGE: 2669 .. 3350

10

(ix) MERKMAL

- (A) NAME/SCHLÜSSEL: 5'UTR
- (B) LAGE: 1 .. 43

																55
CGGAAGCTGG GGAGCATGGA CCAGACCCCG CAGCGCTGGC ACC ATG ACG CTC GCC Met Thr Leu Ala -20																
CGC TTC GTG CTA GCC CTG ATG TTA GGG GCG CTC CCC GAA GTG GTC GGC	Arg Phe Val Leu Ala Leu Met Leu Gly Ala Leu Pro Glu Val Val Gly	103														
-15 -10 -5 -1																
TTT GAT TCT GTC CTC AAT GAT TCC CTC CAC CAC AGC CAC CGC CAT TCG	Phe Asp Ser Val Leu Asn Asp Ser Leu His His Ser His Arg His Ser	151														
1 5 10 15																
CCC CCT GCG GGT CCG CAC TAC CCC TAT TAC CTT CCC ACC CAG CAG CGG	Pro Pro Ala Gly Pro His Tyr Pro Tyr Tyr Leu Pro Thr Gln Gln Arg	199														
20 25 30																
CCC CCG ACG ACG CGT CCG CCG CCG CCT CTC CCG CGC TTC CCG CGC CCC	Pro Pro Thr Thr Arg Pro Pro Pro Pro Leu Pro Arg Phe Pro Arg Pro	247														
35 40 45																
CCG CGG GCG CTC CCT GCC CAG CGC CCG CAC GCC CTC CAG GCC GGG CAC	Pro Arg Ala Leu Pro Ala Gln Arg Pro His Ala Leu Gln Ala Gly His	295														
50 55 60																
ACG CCC CGG CCG CAC CCC TGG GGC TGC CCC GCC GGC GAG CCA TGG GTC	Thr Pro Arg Pro His Pro Trp Gly Cys Pro Ala Gly Glu Pro Trp Val	343														
65 70 75 80																
AGC GTG ACG GAC TTC GGC GCC CCG TGT CTG CGG TGG GCG GAG GTG CCA	Ser Val Thr Asp Phe Gly Ala Pro Cys Leu Arg Trp Ala Glu Val Pro	391														
85 90 95																
CCC TTC CTG GAG CGG TCG CCC CCA GCG AGC TGG GCT CAG CTG CGA GGA	Pro Phe Leu Glu Arg Ser Pro Pro Ala Ser Trp Ala Gln Leu Arg Gly	439														
100 105 110																
CAG CGC CAC AAC TTT TGT CGG AGC CCC GAC GGC GCG GGC AGA CCC TGG	Gln Arg His Asn Phe Cys Arg Ser Pro Asp Gly Ala Gly Arg Pro Trp	487														
115 120 125																
GT TTC TAC GGA GAC GCC CGT GGC AAG GTG GAC TGG GGC TAC TGC GAC	lys Phe Tyr Gly Asp Ala Arg Gly Lys Val Asp Trp Gly Tyr Cys Asp	535														
130 135 140																
GC AGA CAC GGA TCA GTA CGA CTT CGT GGC GGC AAA AAT GAG TTT GAA	ys Arg His Gly Ser Val Arg Leu Arg Gly Gly Lys Asn Glu Phe Glu	583														
45 150 155 160																
GC ACA GTG GAA GTA TAT GCA AGT GGA GTT TGG GGC ACT GTC TGT AGC	ly Thr Val Glu Val Tyr Ala Ser Gly Val Trp Gly Thr Val Cys Ser	631														
165 170 175																
GC CAC TGG GAT GAT TCT GAT GCA TCA GTC ATT TGT CAC CAG CTG CAG	er His Trp Asp Asp Ser Asp Ala Ser Val Ile Cys His Gln Leu Gln	679														
180 185 190																



CTG	GGA	GGA	AAA	GGA	ATA	GCA	AAA	CAA	ACC	CCG	TTT	TCT	GGA	CTG	GGC	727
Leu	Gly	Gly	Lys	Gly	Ile	Ala	Lys	Gln	Thr	Pro	Phe	Ser	Gly	Leu	Gly	
		195					200					205				
CTT	ATT	CCC	ATT	TAT	TGG	AGC	AAT	GTC	CGT	TGC	CGA	GGA	GAT	GAA	GAA	775
Leu	Ile	Pro	Ile	Tyr	Trp	Ser	Asn	Val	Arg	Cys	Arg	Gly	Asp	Glu	Glu	
	210					215					220					
AAT	ATA	CTG	CTT	TGT	GAA	AAA	GAC	ATC	TGG	CAG	GGT	GGG	GTG	TGT	CCT	823
Asn	Ile	Leu	Leu	Cys	Glu	Lys	Asp	Ile	Trp	Gln	Gly	Gly	Val	Cys	Pro	
225				230						235					240	
CAG	AAG	ATG	GCA	GCT	GCT	GTC	ACG	TGT	AGC	TTT	TCC	CAT	GGC	CCA	ACG	871
Gln	Lys	Met	Ala	Ala	Val	Thr	Cys	Ser	Phe	Ser	His	Gly	Pro	Thr		
			245					250					255			
TTC	CCC	ATC	ATT	CGC	CTT	GCT	GGA	GGC	AGC	AGT	GTG	CAT	GAA	GGC	CGG	919
Phe	Pro	Ile	Ile	Arg	Leu	Ala	Gly	Gly	Ser	Ser	Val	His	Glu	Gly	Arg	
		260					265						270			
GTG	GAG	CTC	TAC	CAT	GCT	GGC	CAG	TGG	GGA	ACC	GTT	TGT	GAT	GAC	CAA	967
Val	Glu	Leu	Tyr	His	Ala	Gly	Gln	Trp	Gly	Thr	Val	Cys	Asp	Asp	Gln	
	275					280						285				
TGG	GAT	GAT	GCC	GAT	GCA	GAA	GTG	ATC	TGC	AGG	CAG	CTG	GGC	CTC	AGT	1015
Trp	Asp	Asp	Ala	Asp	Ala	Glu	Val	Ile	Cys	Arg	Gln	Leu	Gly	Leu	Ser	
	290				295						300					
GGC	ATT	GCC	AAA	GCA	TGG	CAT	CAG	GCA	TAT	TTT	GGG	GAA	GGG	TCT	GGC	1063
Gly	Ile	Ala	Lys	Ala	Trp	His	Gln	Ala	Tyr	Phe	Gly	Glu	Gly	Ser	Gly	
305				310					315					320		
CCA	GTT	ATG	TTG	GAT	GAA	GTA	CGC	TGC	ACT	GGG	AAT	GAG	CTT	TCA	ATT	1111
Pro	Val	Met	Leu	Asp	Glu	Val	Arg	Cys	Thr	Gly	Asn	Glu	Leu	Ser	Ile	
			325					330						335		
GAG	CAG	TGT	CCA	AAG	AGC	TCC	TGG	GGA	GAG	CAT	AAC	TGT	GGC	CAT	AAA	1159
Glu	Gln	Cys	Pro	Lys	Ser	Ser	Trp	Gly	Glu	His	Asn	Cys	Gly	His	Lys	
			340					345					350			
GAA	GAT	GCT	GGA	GTG	TCC	TGT	ACC	CCT	CTA	ACA	GAT	GGG	GTC	ATC	AGA	1207
Glu	Asp	Ala	Gly	Val	Ser	Cys	Thr	Pro	Leu	Thr	Asp	Gly	Val	Ile	Arg	
	355					360						365				
CTT	GCA	GGT	GGG	AAA	GGC	AGC	CAT	GAG	GGT	CGC	TTG	GAG	GTA	TAT	TAC	1255
Leu	Ala	Gly	Gly	Lys	Gly	Ser	His	Glu	Gly	Arg	Leu	Glu	Val	Tyr	Tyr	
	370					375					380					
AGA	GGC	CAG	TGG	GGA	ACT	GTC	TGT	GAT	GAT	GGC	TGG	ACT	GAG	CTG	AAT	1303
Arg	Gly	Gln	Trp	Gly	Thr	Val	Cys	Asp	Asp	Gly	Trp	Thr	Glu	Leu	Asn	
385				390						395					400	
ACA	TAC	GTG	GTT	TGT	CGA	CAG	TTG	GGA	TTT	AAA	TAT	GGT	AAA	CAA	GCA	1351
Thr	Tyr	Val	Val	Cys	Arg	Gln	Leu	Gly	Phe	Lys	Tyr	Gly	Lys	Gln	Ala	
			405					410						415		
TCT	GCC	AAC	CAT	TTT	GAA	GAA	AGC	ACA	GGG	CCC	ATA	TGG	TTG	GAT	GAC	1399
Ser	Ala	Asn	His	Phe	Glu	Glu	Ser	Thr	Gly	Pro	Ile	Trp	Leu	Asp	Asp	
			420					425						430		

000007

- 22 -

GTC AGC TGC TCA GGA AAG GAA ACC AGA TTT CTT CAG TGT TCC AGG CGA Val Ser Cys Ser Gly Lys Glu Thr Arg Phe Leu Gln Cys Ser Arg Arg 435 440 445	1447
CAG TGG GGA AGG CAT GAC TGC AGC CAC CGC GAA GAT GTT AGC ATT GCC Gln Trp Gly Arg His Asp Cys Ser His Arg Glu Asp Val Ser Ile Ala 450 455 460	1495
TGC TAC CCT GGC GGC GAG GGA CAC AGG CTC TCT CTG GGT TTT CCT GTC Cys Tyr Pro Gly Gly Glu Gly His Arg Leu Ser Leu Gly Phe Pro Val 465 470 475 480	1543
AGA CTG ATG GAT GGA GAA AAT AAG AAA GAA GGA CGA GTG GAG GTT TTT Arg Leu Met Asp Gly Glu Asn Lys Lys Glu Gly Arg Val Glu Val Phe 485 490 495	1591
ATC AAT GGC CAG TGG GGA ACA ATC TGT GAT GAT GGA TGG ACT GAT AAG Ile Asn Gly Gln Trp Gly Thr Ile Cys Asp Asp Gly Trp Thr Asp Lys 500 505 510	1639
GAT GCA GCT GTG ATC TGT CGT CAG CTT GGC TAC AAG GGT CCT GCC AGA Asp Ala Ala Val Ile Cys Arg Gln Leu Gly Tyr Lys Gly Pro Ala Arg 515 520 525	1687
GCA AGA ACC ATG GCT TAC TTT GGA GAA GGA AAA GGA CCC ATC CAT GTG Ala Arg Thr Met Ala Tyr Phe Gly Glu Gly Lys Gly Pro Ile His Val 530 535 540	1735
GAT AAT GTG AAG TGC ACA GGA AAT GAG AGG TCC TTG GCT GAC TGT ATC Asp Asn Val Lys Cys Thr Gly Asn Glu Arg Ser Leu Ala Asp Cys Ile 545 550 555 560	1783
AAG CAA GAT ATT GGA AGA CAC AAC TGC CGC CAC AGT GAA GAT GCA GGA Lys Gln Asp Ile Gly Arg His Asn Cys Arg His Ser Glu Asp Ala Gly 565 570 575	1831
GTT ATT TGT GAT TAT TTT GGC AAG AAG GCC TCA GGT AAC AGT AAT AAA Val Ile Cys Asp Tyr Phe Gly Lys Lys Ala Ser Gly Asn Ser Asn Lys 580 585 590	1879
GAG TCC CTC TCA TCT GTT TGT GGC TTG AGA TTA CTG CAC CGT CGG CAG Glu Ser Leu Ser Ser Val Cys Gly Leu Arg Leu Leu His Arg Arg Gln 595 600 605	1927
AAG CGG ATC ATT GGT GGG AAA AAT TCT TTA AGG GGT GGT TGG CCT TGG Lys Arg Ile Ile Gly Gly Lys Asn Ser Leu Arg Gly Gly Trp Pro Trp 610 615 620	1975
CAG GTT TCC CTC CGG CTG AAG TCA TCC CAT GGA GAT GGC AGG CTC CTC Gln Val Ser Leu Arg Leu Lys Ser Ser His Gly Asp Gly Arg Leu Leu 625 630 635 640	2023
TGC GGG GCT ACG CTC CTG AGT AGC TGC TGG GTC CTC ACA GCA GCA CAC Cys Gly Ala Thr Leu Leu Ser Ser Cys Trp Val Leu Thr Ala Ala His 645 650 655	2071
TGT TTC AAG AGG TAT GGC AAC AGC ACT AGG AGC TAT GCT GTT AGG GTT Cys Phe Lys Arg Tyr Gly Asn Ser Thr Arg Ser Tyr Ala Val Arg Val 660 665 670	2119

- 23 -

GGA GAT TAT CAT ACT CTG GTA CCA GAG GAG TTT GAG GAA GAA ATT GGA	2167
Gly Asp Tyr His Thr Leu Val Pro Glu Glu Phe Glu Glu Glu Ile Gly	
675 680 685	
GTT CAA CAG ATT GTG ATT CAT CGG GAG TAT CGA CCC GAC CGC AGT GAT	2215
Val Gln Gln Ile Val Ile His Arg Glu Tyr Arg Pro Asp Arg Ser Asp	
690 695 700	
TAT GAC ATA GCC CTG GTT AGA TTA CAA GGA CCA GAA GAG CAA TGT GCC	2263
Tyr Asp Ile Ala Leu Val Arg Leu Gln Gly Pro Glu Glu Gln Cys Ala	
705 710 715 720	
AGA TTC AGC AGC CAT GTT TTG CCA GCC TGT TTA CCA CTC TGG AGA GAG	2311
Arg Phe Ser Ser His Val Leu Pro Ala Cys Leu Pro Leu Trp Arg Glu	
725 730 735	
AGG CCA CAG AAA ACA GCA TCC AAC TGT TAC ATA ACA GGA TGG GGT GAC	2359
Arg Pro Gln Lys Thr Ala Ser Asn Cys Tyr Ile Thr Gly Trp Gly Asp	
740 745 750	
ACA GGA CGA GCC TAT TCA AGA ACA CTA CAA CAA GCA GCC ATT CCC TTA	2407
Thr Gly Arg Ala Tyr Ser Arg Thr Leu Gln Gln Ala Ala Ile Pro Leu	
755 760 765	
CTT CCT AAA AGG TTT TGT GAA GAA CGT TAT AAG GGT CGG TTT ACA GGG	2455
Leu Pro Lys Arg Phe Cys Glu Glu Arg Tyr Lys Gly Arg Phe Thr Gly	
770 775 780	
AGA ATG CTT TGT GCT GGA AAC CTC CAT GAA CAC AAA CGC GTG GAC AGC	2503
Arg Met Leu Cys Ala Gly Asn Leu His Glu His Lys Arg Val Asp Ser	
785 790 795 800	
TGC CAG GGA GAC AGC GGA GGA CCA CTC ATG TGT GAA CGG CCC GGA GAG	2551
Cys Gln Gly Asp Ser Gly Gly Pro Leu Met Cys Glu Arg Pro Gly Glu	
805 810 815	
AGC TGG GTG GTG TAT GGG GTG ACC TCC TGG GGG TAT GGC TGT GGA GTC	2599
Ser Trp Val Val Tyr Gly Val Thr Ser Trp Gly Tyr Gly Cys Gly Val	
820 825 830	
AAG GAT TCT CCT GGT GTT TAT ACC AAA GTC TCA GCC TTT GTA CCT TGG	2647
Lys Asp Ser Pro Gly Val Tyr Thr Lys Val Ser Ala Phe Val Pro Trp	
835 840 845	
ATA AAA AGT GTC ACC AAA CTG TAA TTCTTCATGG AAACCTCAAA GCAGCATTT	2700
Ile Lys Ser Val Thr Lys Leu *	
850 855	
AAACAAATGG AAAACTTTGA ACCCCCACTA TTAGCACTCA GCAGAGATGA CAACAAATGG	2760
CAAGATCTGT TTTTGCTTTG TGTGTGGTA AAAAATTGTG TACCCCTGCTG TGCTTTTGAG	2820
AAATTTGTGA ACATTTTCAG AGGCCTCAGT GTAGTGGAAG TGATAATCCT TAAATGAACA	2880
TTTTCTACCC TAATTTCACT GGAGTGA CTATTCTAAGCC TCATCTATCC CCTACCTATT	2940

988.97

- 24 -

TCTCAAAATC ATTCTATGCT GATTTTACAA AAGATCATTT TTACATTTGA ACTGAGAACC 3000  
CCTTTTAATT GAATCAGTGG TGTCTGAAAT CATATTAAAT ACCCACATTT GACATAAATG 3060  
CGGTACCCTT TACTACACTC ATGAGTGGCA TATTTATGCT TAGGTCTTTT CAAAAGACTT 3120  
GACAAGAAAT CTTCATATTC TCTGTAGCCT TTGTCAAGTG AGGAAATCAG TGGTTAAAGA 3180  
ATTCCACTAT AAACTTT TAG GCCTGAATAG GAGTAGTAAA GCCTCAAGGA CATCTGCCTG 3240  
TCACAATATA TTCTCAAAGT GATCTGATAT TTGGAAACAA GTATCCTTGT TGAGTACCAA 3300  
GTGCTACAGA AACCATAAGA TAAAATACT TTCTACCTAC AGCGTGCCCCG 3350

(1) ANGABEN ZUR VERBINDUNG DER FORMEL II (Neurotrypsin der Maus)

## (i) SEQUENZKENNZEICHEN:

- 5 (A) LÄNGE: 2376 Basenpaare  
(B) ART: Nucleinsäure  
(C) STRANGFORM: Einzelstrang  
(D) TOPOLOGIE: linear

## 10 (ii) ART DES MOLEKÜLS: c-DNA zu m-RNA

## (vi) URSPRÜNGLICHE HERKUNFT:

- (A) ORGANISMUS: Mus musculus  
15 (D) ENTWICKLUNGSSTADIUM: Postnatal-Tag 10  
(F) GEWEBETYP: Gehirn  
(G) ZELLTYP: Neuronen

## (vii) UNMITTELBARE HERKUNFT:

20

- (A) BIBLIOTHEK: mouse brain cDNA library in the lambda Uni-ZAP-XR vector,  
oligo (dT)-primed, from Balb c mice, postnatal day 20,  
Cat. No.. 937 319; Stratagene, La Jolla, CA, USA

25

- (B) CLONE: cDNA clone no. 16

## (vii) UNMITTELBARE HERKUNFT:

- 30 (A) BIBLIOTHEK: mouse brain cDNA library in the Lambda gt10 vector,  
oligo(dT)- and random-primed, embryonic day 15,  
Cat. No. ML 3002a; Clontech, Palo Alto, CA, USA

986/97

- 26 -

(B) CLONE: cDNA clone #25

(ix) MERKMAL:

5

(A) NAME/SCHLÜSSEL: Signalpeptid

(B) LAGE: 24 .. 86

10 (ix) MERKMAL:

(A) NAME/SCHLÜSSEL: reifes Peptid

(B) LAGE: 87 .. 2306

15

(ix) MERKMAL:

(A) NAME/SCHLÜSSEL: codierende Sequenz

(B) LAGE: 24 .. 2306

20

(ix) MERKMAL:

(A) NAME/SCHLÜSSEL: Protein-reiches basisches Segment

25 (B) LAGE: 90 .. 275

(ix) MERKMAL:

30 (A) NAME/SCHLÜSSEL: Kringle-Domäne

(B) LAGE: 276 .. 494

(ix) MERKMAL:

(A) NAME/SCHLÜSSEL: SRCR-Domäne 1

5 (B) LAGE: 519 .. 824

(ix) MERKMAL:

10 (A) NAME/SCHLÜSSEL: SRCR-Domäne 2

(B) LAGE: 840 .. 1142

(ix) MERKMAL:

15

(A) NAME/SCHLÜSSEL: SRCR-Domäne 3

(B) LAGE: 1179 .. 1484

20 (ix) MERKMAL:

(A) NAME/SCHLÜSSEL: proteolytische Domäne

(B) LAGE: 1536 .. 2306

25

(ix) MERKMAL:

(A) NAME/SCHLÜSSEL: Histidin der katalytischen Triade

(B) LAGE: 1707 .. 1709

30

986/97

(ix) MERKMAL:

- 5 (A) NAME/SCHLÜSSEL: Asparaginsäure der katalytischen Triade  
(B) LAGE: 1857 .. 1859

(ix) MERKMAL:

- 10 (A) NAME/SCHLÜSSEL: Serin der katalytischen Triade  
(B) LAGE: 2154 .. 2156

(ix) MERKMAL:

- 15 (A) NAME/SCHLÜSSEL: PolyA-Signal  
(B) LAGE: 2324 .. 2329 und 2331 .. 2336

(ix) MERKMAL:

- 20 (A) NAME/SCHLÜSSEL: PolyA-Bereich  
(B) LAGE: 2357 .. 2376

(ix) MERKMAL:

25

- (A) NAME/SCHLÜSSEL: 3'UTR  
(B) LAGE: 2307 .. 2341 oder 2307 .. 2356



906/97

- 29 -

(ix) MERKMAL:

(A) NAME/SCHLÜSSEL: 5'UTR

5 (B) LAGE: 1 .. 23

005.97

- 30 -

# Verbindung der Formel II (Neurotrypsin der Maus)

GGACCACACT	CGGCGCCGCA	GCC	ATG	GCG	CTC	GCC	CGC	TGC	GTG	CTG	GCT	GTG	53
			Met	Ala	Leu	Ala	Arg	Cys	Val	Leu	Ala	Val	
			-20						-15				
ATT	TTA	GGG	GCA	CTG	TCT	GTA	GTG	GCC	CGC	GCT	GAT	CCG	101
Ile	Leu	Gly	Ala	Leu	Ser	Val	Ala	Arg	Ala	Asp	Pro	Val	
-10					-5					1		5	
TCT	CCC	CTT	CAC	CGC	CCG	CAT	CCG	TCC	CCA	CCG	CGT	TCC	149
Ser	Pro	Leu	His	Arg	Pro	His	Pro	Ser	Pro	Pro	Arg	Ser	
			10					15				20	
CAC	TAC	CTT	CCC	AGC	TCG	CGG	CGG	CCA	CCC	AGG	ACC	CCG	197
His	Tyr	Leu	Pro	Ser	Arg	Arg	Pro	Pro	Pro	Arg	Thr	Pro	
			25				30					35	
CTC	CCG	CTG	CGG	ATC	CCC	GCT	GCC	CAG	CGC	CCG	CAG	GTC	245
Leu	Pro	Leu	Arg	Ile	Pro	Ala	Ala	Gln	Arg	Pro	Gln	Val	
	40					45					50		
GGG	CAC	ACG	CCC	CCG	ACG	ATT	CCA	CGC	CGC	TGC	GGG	GCA	293
Gly	His	Thr	Pro	Pro	Thr	Ile	Pro	Arg	Arg	Cys	Gly	Ala	
55						60				65			
TGG	GGC	AAT	GCC	ACC	AAC	CTC	GGC	GTC	CCG	TGT	CTA	CAC	341
Trp	Gly	Asn	Ala	Thr	Asn	Leu	Gly	Val	Pro	Cys	Leu	His	
70					75				80			85	
GTG	CCG	CCC	TTC	CTG	GAG	CGG	TCG	CCC	CCG	GCC	AGT	TGG	389
Val	Pro	Pro	Phe	Leu	Glu	Arg	Ser	Pro	Pro	Ala	Ser	Trp	
				90				95				100	
CGA	GGG	CAG	CCG	CAC	AAC	TTC	TGC	CGG	AGC	CCG	GAT	GGC	437
Arg	Gly	Gln	Pro	His	Asn	Phe	Cys	Arg	Ser	Pro	Asp	Gly	
			105				110					115	
CCT	TGG	TGC	TTC	TAT	CGG	AAT	GCC	CAG	GGC	AAA	GTA	GAC	485
Pro	Trp	Cys	Phe	Tyr	Arg	Asn	Ala	Gln	Gly	Lys	Val	Asp	
		120					125					130	
TGC	GAT	TGT	GGT	CAA	GGC	CCG	GCG	TTG	CCC	GTC	ATT	CGC	533
Cys	Asp	Cys	Gly	Gln	Gly	Pro	Ala	Leu	Pro	Val	Ile	Arg	
135						140					145		
GGG	AAC	AGT	GGG	CAT	GAA	GGT	CGA	GTG	GAG	CTG	TAC	CAC	581
Gly	Asn	Ser	Gly	His	Glu	Gly	Arg	Val	Glu	Leu	Tyr	His	
150					155				160			165	
TGG	GGG	ACC	ATC	TGT	GAC	GAC	CAA	TGG	GAC	AAT	GCA	GAC	629
Trp	Gly	Thr	Ile	Cys	Asp	Asp	Gln	Trp	Asp	Asn	Ala	Asp	
				170					175			180	
ATC	TGT	AGG	CAG	CTG	GGG	CTC	AGT	GGC	ATT	GCC	AAA	GCA	677
Ile	Cys	Arg	Gln	Leu	Gly	Leu	Ser	Gly	Ile	Ala	Lys	Ala	
			185					190				195	

GCA CAT TTT GGG GAA GGA TCT GGC CCA ATA TTG TTG GAT GAA GTA CGC Ala His Phe Gly Glu Gly Ser Gly Pro Ile Leu Leu Asp Glu Val Arg 200 205 210	725
TGC ACC GGA AAC GAG CTG TCA ATT GAG CAA TGT CCA AAG AGT TCC TGG Cys Thr Gly Asn Glu Leu Ser Ile Glu Gln Cys Pro Lys Ser Ser Trp 215 220 225	773
GGC GAA CAT AAC TGT GGC CAT AAA GAA GAT GCT GGA GTG TCT TGT GTT Gly Glu His Asn Cys Gly His Lys Glu Asp Ala Gly Val Ser Cys Val 230 235 240 245	821
CCT CTA ACA GAT GGT GTC ATC AGA CTG GCA GGA GGA AAA AGT ACC CAT Pro Leu Thr Asp Gly Val Ile Arg Leu Ala Gly Gly Lys Ser Thr His 250 255 260	869
GAA GGT CGC CTG GAG GTC TAC TAC AAG GGG CAG TGG GGG ACA GTC TGT Glu Gly Arg Leu Glu Val Tyr Tyr Lys Gly Gln Trp Gly Thr Val Cys 265 270 275	917
GAT GAT GGC TGG ACT GAG ATG AAC ACA TAC GTG GCT TGT CGA CTG CTG Asp Asp Gly Trp Thr Glu Met Asn Thr Tyr Val Ala Cys Arg Leu Leu 280 285 290	965
GGA TTT AAA TAC GGC AAA CAG TCC TCT GTG AAC CAT TTT GAT GGC AGC Gly Phe Lys Tyr Gly Lys Gln Ser Ser Val Asn His Phe Asp Gly Ser 295 300 305	1013
AAC AGG CCC ATA TGG CTG GAT GAC GTC AGC TGC TCA GGA AAA GAA GTC Asn Arg Pro Ile Trp Leu Asp Asp Val Ser Cys Ser Gly Lys Glu Val 310 315 320 325	1061
AGC TTC ATT CAG TGT TCC AGG AGA CAG TGG GGA AGG CAT GAC TGC AGC Ser Phe Ile Gln Cys Ser Arg Arg Gln Trp Gly Arg His Asp Cys Ser 330 335 340	1109
CAT AGA GAA GAT GTG GGC CTC ACC TGC TAT CCT GAC AGC GAT GGA CAT His Arg Glu Asp Val Gly Leu Thr Cys Tyr Pro Asp Ser Asp Gly His 345 350 355	1157
AGG CTT TCT CCA GGT TTT CCC ATC AGA CTA GTG GAT GGA GAG AAT AAG Arg Leu Ser Pro Gly Phe Pro Ile Arg Leu Val Asp Gly Glu Asn Lys 360 365 370	1205
AAG GAA GGA CGA GTG GAG GTT TTT GTC AAT GGC CAA TGG GGA ACA ATC Lys Glu Gly Arg Val Glu Val Phe Val Asn Gly Gln Trp Gly Thr Ile 375 380 385	1253
TGC GAT GAC GGA TGG ACC GAT AAG CAT GCA GCT GTG ATC TGC CGG CAA Cys Asp Asp Gly Trp Thr Asp Lys His Ala Ala Val Ile Cys Arg Gln 390 395 400 405	1301
CTT GGC TAT AAG GGT CCT GCC AGA GCA AGG ACT ATG GCT TAT TTT GGG Leu Gly Tyr Lys Gly Pro Ala Arg Ala Arg Thr Met Ala Tyr Phe Gly 410 415 420	1349
GAA GGA AAA GGC CCC ATC CAC ATG GAT AAT GTG AAG TGC ACA GGA AAT Glu Gly Lys Gly Pro Ile His Met Asp Asn Val Lys Cys Thr Gly Asn 425 430 435	1397

GAG Glu	AAG Lys	GCC Ala	CTG Leu	GCT Ala	GAC Asp	TGT Cys	GTC Val	AAA Lys	CAA Gln	GAC Asp	ATT Ile	GGA Gly	AGG Arg	CAC His	AAC Asn	1445
		440					445					450				
TGC Cys	CGC Arg	CAC His	AGT Ser	GAG Glu	GAT Asp	GCA Ala	GGA Gly	GTC Val	ATC Ile	TGT Cys	GAC Asp	TAT Tyr	TTA Leu	GAG Glu	AAG Lys	1493
	455					460					465					
AAA Lys	GCA Ala	TCA Ser	AGT Ser	AGT Ser	GGT Gly	AAT Asn	AAA Lys	GAG Glu	ATG Met	CTC Leu	TCA Ser	TCT Ser	GGA Gly	TGT Cys	GGA Gly	1541
470					475					480					485	
CTG Leu	AGG Arg	TTA Leu	CTG Leu	CAC His	CGT Arg	CGG Arg	CAG Gln	AAA Lys	CGG Arg	ATC Ile	ATT Ile	GGT Gly	GGG Gly	AAC Asn	AAT Asn	1589
				490					495					500		
TCT Ser	TTA Leu	AGG Arg	GGT Gly	GCC Ala	TGG Trp	CCT Pro	TGG Trp	CAG Gln	GCT Ala	TCC Ser	CTC Leu	AGG Arg	CTG Leu	AGG Arg	TCG Ser	1637
			505					510					515			
GCC Ala	CAT His	GGA Gly	GAC Asp	GGC Gly	AGG Arg	CTG Leu	CTT Leu	TGT Cys	GGA Gly	GCT Ala	ACC Thr	CTT Leu	CTG Leu	AGT Ser	AGC Ser	1685
		520					525					530				
TGC Cys	TGG Trp	GTC Val	CTG Leu	ACA Thr	GCT Ala	GCA Ala	CAC His	TGC Cys	TTC Phe	AAA Lys	AGG Arg	TAC Tyr	GGA Gly	AAC Asn	AAC Asn	1733
	535					540					545					
TCG Ser	AGG Arg	AGC Ser	TAT Tyr	GCA Ala	GTT Val	CGA Arg	GTT Val	GGG Gly	GAT Asp	TAT Tyr	CAT His	ACT Thr	CTG Leu	GTC Val	CCA Pro	1781
550					555					560					565	
GAG Glu	GAG Glu	TTT Phe	GAA Glu	CAA Gln	GAA Glu	ATA Ile	GGG Gly	GTT Val	CAA Gln	CAG Gln	ATT Ile	GTG Val	ATT Ile	CAC His	AGG Arg	1829
				570					575					580		
AAC Asn	TAC Tyr	AGG Arg	CCA Pro	GAC Asp	AGA Arg	AGC Ser	GAC Asp	TAT Tyr	GAC Asp	ATT Ile	GCC Ala	CTG Leu	GTT Val	AGA Arg	TTG Leu	1877
			585					590					595			
CAA Gln	GGA Gly	CCA Pro	GGG Gly	GAG Glu	CAA Gln	TGT Cys	GCC Ala	AGA Arg	CTA Leu	AGC Ser	ACC Thr	CAC His	GTT Val	TTG Leu	CCA Pro	1925
		600					605					610				
GCC Ala	TGT Cys	TTA Leu	CCT Pro	CTA Leu	TGG Trp	AGA Arg	GAG Glu	AGG Arg	CCA Pro	CAG Gln	AAA Lys	ACA Thr	GCC Ala	TCC Ser	AAC Asn	1973
	615					620					625					
TGT Cys	CAC His	ATA Ile	ACA Thr	GGA Gly	TGG Trp	GGA Gly	GAC Asp	ACA Thr	GGT Gly	CGT Arg	GCC Ala	TAC Tyr	TCA Ser	AGA Arg	ACT Thr	2021
630					635					640					645	
CTA Leu	CAA Gln	CAA Gln	GCT Ala	GCT Ala	GTG Val	CCT Pro	CTG Leu	TTA Leu	CCC Pro	AAG Lys	AGG Arg	TTT Phe	TGT Cys	AAA Lys	GAG Glu	2069
				650					655					660		
AGG Arg	TAC Tyr	AAG Lys	GGA Gly	CTA Leu	TTT Phe	ACT Thr	GGG Gly	AGA Arg	ATG Met	CTC Leu	TGT Cys	GCT Ala	GGG Gly	AAC Asn	CTC Leu	2117
			665					670					675			

CAA GAA GAC AAC CGT GTG GAC AGC TGC CAG GGA GAC AGT GGA GGA CCA	2165
Gln Glu Asp Asn Arg Val Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro	
680 685 690	
CTC ATG TGT GAA AAG CCT GAT GAG TCC TGG GTT GTG TAT GGG GTG ACT	2213
Leu Met Cys Glu Lys Pro Asp Glu Ser Trp Val Val Tyr Gly Val Thr	
695 700 705	
TCC TGG GGG TAT GGA TGT GGA GTC AAA GAC ACT CCT GGA GTT TAT ACC	2261
Ser Trp Gly Tyr Gly Cys Gly Val Lys Asp Thr Pro Gly Val Tyr Thr	
710 715 720 725	
AGA GTC CCC GCT TTT GTA CCT TGG ATA AAA AGT GTC ACC AGT CTG	2306
Arg Val Pro Ala Phe Val Pro Trp Ile Lys Ser Val Thr Ser Leu	
730 735 740	
TAAGTTATGG AAAGCTCAAG AAATAGTAAA ACAGTAACTA TTCAGTCTTC AAAAAAAAAA	2366
AAAAAAAAAA	2376

**Patentansprüche****1. Neurotrypsine der Formeln I und II**

5

**I: Neurotrypsin des Menschen****II: Neurotrypsin der Maus**

10 einschliesslich der separaten, codierenden und codierten Sequenzen dieser Verbindungen der Formeln I oder II, einschliesslich der separaten Teilsequenzen der codierenden und codierten Sequenzen dieser Verbindungen der Formeln I oder II, wie zum Beispiel die codierenden und codierten Sequenzen der katalytischen Domänen der Verbindungen der Formeln I oder II, einschliesslich der codierenden oder codierten Sequenzen oder Teilsequenzen der den Verbindungen  
15 der Formeln I oder II entsprechenden Splice-Varianten, einschliesslich der codierenden oder codierten Sequenzen oder Teilsequenzen der den Verbindungen der Formeln I oder II entsprechenden Allele, einschliesslich aller Sequenzvarianten der codierenden oder codierten Sequenzen, oder Teilen davon, der Verbindungen der Formeln I oder II, deren biologische Aktivität derjenigen der Verbindungen der  
20 Formeln I oder II gleich oder ähnlich ist, beispielsweise Sequenzvarianten der Verbindungen der Formeln I oder II, welche an den nicht konservierten Aminosäuresequenzpositionen von der Sequenz der Formeln I oder II abweichen, einschliesslich der an die codierenden Sequenzen, oder Teile davon, unter stringenten Bedingungen hybridisierenden Sequenzen, einschliesslich der  
25 Translationsprodukte der an die codierenden Sequenzen der Verbindungen der Formeln I oder II, oder an Teile davon, unter stringenten Bedingungen hybridisierenden Sequenzen, einschliesslich der Nucleotidsequenzen, welche die durch die Verbindungen der Formeln I oder II codierten Proteine, oder Teile davon, codieren, aber, als Resultat unterschiedlicher Verwendung alternativer Codons,  
30 degeneriert sind bezüglich der als Verbindungen der Formel I oder II definierten Nucleotidsequenzen.

2. Medikament, dadurch gekennzeichnet, dass es als wenigstens eine

aktive Verbindung entweder die codierte Sequenz oder die codierende Sequenz der Verbindung der Formel I oder der Formel II, oder die separaten Teilsequenzen der codierten und codierenden Sequenzen dieser Verbindungen der Formel I oder II, wie zum Beispiel die codierenden oder codierten Sequenzen der katalytischen Domänen, enthält, einschliesslich der codierenden oder codierten Sequenzen oder Teilsequenzen der den Verbindungen der Formeln I oder II entsprechenden Splice-Varianten, einschliesslich der codierenden oder codierten Sequenzen oder Teilsequenzen der den Verbindungen der Formeln I oder II entsprechenden Allele, einschliesslich aller Sequenzvarianten, der codierenden oder codierten Sequenzen, oder Teilen davon, der Verbindungen der Formeln I oder II, deren biologische Aktivität derjenigen der Verbindungen der Formeln I oder II gleich oder ähnlich ist, beispielsweise Sequenzvarianten der Verbindungen der Formeln I oder II, welche an den nicht konservierten Aminosäuresequenzpositionen von der Sequenz der Formeln I oder II abweichen, einschliesslich der an die codierenden Sequenzen, oder Teile davon, unter stringenten Bedingungen hybridisierenden Sequenzen, einschliesslich der Translationsprodukte der an die codierenden Sequenzen der Verbindungen der Formeln I oder II, oder an Teile davon, unter stringenten Bedingungen hybridisierenden Sequenzen, einschliesslich der Nucleotidsequenzen, welche die durch die Verbindungen der Formeln I oder II codierten Proteine, oder Teile davon, codieren, aber, als Resultat unterschiedlicher Verwendung alternativer Codons, degeneriert sind bezüglich der als Verbindungen der Formel I oder II definierten Nucleotidsequenzen.

3. Medikament, dadurch gekennzeichnet, dass es als wenigstens eine aktive Verbindung eine Substanz enthält, welche die Funktion der codierten Sequenz der Verbindungen der Formeln I oder II verändert, zum Beispiel, indem es die katalytische Wirkung des codierten Proteins, oder eines Teils davon, vermindert oder verstärkt, oder indem es die Verweildauer des codierten Proteins an dessen Wirkungsort im Körper verkürzt oder verlängert.

30

4. Medikament, dadurch gekennzeichnet, dass es als wenigstens eine aktive Verbindung eine Substanz enthält, welche die Expression der codierenden oder codierten Sequenzen der Verbindungen der Formel I oder II verändert, zum

Beispiel indem es die Transkription der mRNA fördert oder hemmt, oder indem es die Translation der codierten Sequenzen der Formeln I oder II fördert oder hemmt.

5 5. Medikament nach Anspruch 2, 3 oder 4, dadurch gekennzeichnet, dass es die Apoptose von Zellen des Nervensystems verhindert.

6. Medikament nach Anspruch 5, dadurch gekennzeichnet, dass es sich bei der Apoptose um Apoptose im Zusammenhang mit Schädigungen des Nervengewebes handelt, wie beispielsweise Gehirnfarkt, oder Gehirnblutung, oder  
10 Gehirntrauma.

7. Medikament nach Anspruch 5, dadurch gekennzeichnet, dass es sich bei der Apoptose um Apoptose im Zusammenhang mit Schädigungen des Nervengewebes handelt, welche auf Grund von Sauerstoffmangel oder von  
15 Vergiftungen auftreten.

8. Medikament nach Anspruch 2, 3 oder 4, dadurch gekennzeichnet, dass es die Regeneration von verletztem, beschädigtem, unterentwickeltem, oder fehlentwickeltem Gehirn- und/oder Nervengewebe beeinflusst.  
20

9. Medikament nach Anspruch 2, 3 oder 4, dadurch gekennzeichnet, dass es nach Gehirn- und/oder Nervenverletzungen oder nach der Zerstörung oder Beschädigung von Gehirnarealen die Reorganisation von intakt gebliebenen Gehirn-, respektive Nervenarealen fördert.  
25

10. Medikament nach Anspruch 2, 3 oder 4, dadurch gekennzeichnet, dass es pathologische Schmerzzustände verhindert, lindert oder behebt.

11. Medikament nach Anspruch 2, 3 oder 4, dadurch gekennzeichnet, dass  
30 es zur Förderung der Gehirnleistung bei gesunden Personen, sowie bei Personen mit reduzierter Gehirnleistung beiträgt.

12. Medikament nach Anspruch 2, 3 oder 4, dadurch gekennzeichnet, dass



es Lern- und Gedächtnisfunktionen bei gesunden Personen, sowie bei Personen mit reduzierten Lern- und Gedächtnisfunktionen verbessert.

5 13. Medikament nach Anspruch 2, 3 oder 4, dadurch gekennzeichnet, dass es Störungen aus dem Formenkreis der Störungen des psychischen Wohlbefindens, oder der psychosomatischen Befindlichkeit, wie beispielsweise Nervosität oder "Innere Unruhe", lindert oder behebt.

10 14. Medikament nach Anspruch 2, 3 oder 4, dadurch gekennzeichnet, dass es Störungen im Bereich der emotionellen Funktionen, wie zum Beispiel Angstzustände, lindert oder behebt.

15 15. Medikament nach Anspruch 2, 3 oder 4, dadurch gekennzeichnet, dass es psychiatrische Störungen lindert oder behebt.

20 16. Medikament nach Anspruch 15, dadurch gekennzeichnet, dass es sich um eine Störung aus dem Formenkreis der Schizophrenie und schizophrenieartiger Störungen handelt, einschliesslich chronischer Schizophrenie, chronischer schizoaffektiver Störungen, unspezifischer Störungen, einschliesslich akuter und chronischer Schizophrenie verschiedener Ausprägung, wie beispielsweise schwere, nicht-remittierende "Kraepelinsche" Schizophrenie, oder wie beispielsweise der DSM-III-R-Prototyp der schizophrenieartigen Störungen, einschliesslich episodischer schizophrener Störungen, einschliesslich wahnhafter schizophrenieartiger Störungen, einschliesslich Schizophrenie-ähnlicher Persönlichkeitsstörungen, wie beispielsweise  
25 schizophrenieartiger Persönlichkeitsstörungen mit milderer Symptomatik, einschliesslich schizotypischer Persönlichkeitsstörungen, einschliesslich der latenten Formen schizophrener oder schizophrenieartiger Störungen, einschliesslich nicht-organischer psychotischer Störungen.

30 17. Medikament nach Anspruch 15, dadurch gekennzeichnet, dass es sich um Störungen aus dem Formenkreis der endogenen Depressionen oder aus dem Formenkreis der manischen und manisch-depressiven Störungen handelt.

18. Medikament nach Anspruch 2, 3 oder 4, dadurch gekennzeichnet, dass es durch mangelhafte oder überfunktionelle Proteasen bedingte Störungen der Gehirnfunktion lindert oder behebt.

5        19. Medikament nach Anspruch 18, dadurch gekennzeichnet, dass die Protease Gewebe-Plasminogenaktivator, abgekürzt mit tPA, Urokinase-Plasminogenaktivator, abgekürzt mit uPA, oder Plasmin ist.

10       20. Medikament nach Anspruch 2, 3 oder 4, dadurch gekennzeichnet, dass es durch mangelhafte oder überfunktionelle Proteasen bedingte Störungen der Lungenfunktion lindert oder behebt.

15       21. Medikament nach Anspruch 20, dadurch gekennzeichnet, dass es sich bei der Störung der Lungenfunktion um chronische Bronchitis oder Lungenemphysem handelt.

22. Verwendung der codierenden Nucleotidsequenzen der Verbindungen der Formeln I oder II, einschliesslich der separaten Teilsequenzen der codierenden Sequenzen der Verbindungen der Formeln I oder II, wie zum Beispiel die codierenden Sequenzen der katalytischen Domänen der Verbindungen der Formeln I oder II, 20 einschliesslich der codierenden Nucleotidsequenzen oder Teilsequenzen der den Verbindungen der Formeln I oder II entsprechenden Splice-Varianten, einschliesslich der codierenden Sequenzen, oder Teilsequenzen davon, der den Verbindungen der Formeln I oder II entsprechenden Allele, einschliesslich aller Sequenzvarianten der 25 codierenden Sequenzen, oder Teilen davon, der Verbindungen der Formeln I oder II, deren Translationsprodukte bezüglich ihrer biologischen Aktivität den Translationsprodukten der Verbindungen der Formeln I oder II gleich oder ähnlich sind, beispielsweise Sequenzvarianten der Verbindungen der Formeln I oder II, welche an den nicht konservierten Aminosäuresequenzpositionen von der Sequenz 30 der Formeln I oder II abweichen, einschliesslich der an die codierenden Sequenzen der Verbindungen der Formeln I oder II, oder an Teile davon, unter stringenten Bedingungen hybridisierenden Sequenzen, einschliesslich der Nucleotid-Sequenzen, welche die durch die Verbindungen der Formeln I oder II codierten Proteine, oder

Teile davon, codieren, aber, als Resultat unterschiedlicher Verwendung alternativer Codons, degeneriert sind bezüglich der als Verbindungen der Formel I oder II definierten Nucleotidsequenzen, für die Herstellung recombinanter Proteine.

5           23. Verwendung von Proteinen mit den codierten Aminosäuresequenzen der Verbindungen der Formeln I oder II, einschliesslich der Proteine mit den separaten Teilsequenzen der codierten Aminosäuresequenzen der Verbindungen der Formeln I oder II, wie beispielsweise die separaten katalytischen Domänen der Verbindungen der Formeln I oder II, einschliesslich der Proteine mit den codierten Sequenzen oder  
10 Teilsequenzen der den Verbindungen der Formeln I oder II entsprechenden Splice-Varianten, einschliesslich der Proteine mit den codierten Aminosäuresequenzen, oder Teilsequenzen davon, der den Verbindungen der Formeln I oder II entsprechenden Allele, einschliesslich aller Sequenzvarianten der codierten Sequenzen, oder Teilen davon, der Verbindungen der Formeln I oder II, deren biologische Aktivität derjenigen  
15 der codierten Sequenzen der Verbindungen der Formeln I oder II gleich oder ähnlich ist, beispielsweise Sequenzvarianten der Verbindungen der Formeln I oder II, welche an den nicht konservierten Aminosäuresequenzpositionen von der Sequenz der Formeln I oder II abweichen, einschliesslich der Proteine mit den codierten Aminosäuresequenzen, oder Teilsequenzen davon, der an die codierenden  
20 Sequenzen der Verbindungen der Formeln I oder II, oder an Teile davon, unter stringenten Bedingungen hybridisierenden Sequenzen, als Zielstruktur für die Entwicklung von Pharmaca, zum Beispiel zur Hemmung oder Förderung der katalytischen Wirkung der codierten Proteine der Formeln I oder II.

25           24. Verwendung der Spezies-homologen Proteine, oder Teilen davon, der Verbindungen der Formeln I oder II, wie beispielsweise die Spezies-homologen Proteine der Ratte, des Kaninchens, des Rindes, des Schafes, des Schweins, der Primaten, der Vögel, des Zebrafisches, der Taufliege (*Drosophila melanogaster*), etc., einschliesslich der Teilsequenzen davon, wie beispielsweise die separaten  
30 katalytischen Domänen, einschliesslich der Splice-Varianten der Spezies-homologen Proteine, einschliesslich der Allele der Spezies-homologen Proteine, einschliesslich der Translationsprodukte der an die codierenden Sequenzen oder Teilsequenzen der den Verbindungen der Formeln I oder II entsprechenden Spezies-homologen

Verbindungen, oder deren Splice-Varianten, oder deren Allele, unter stringenten Verbindungen hybridisierenden Sequenzen, als Zielstruktur für die Entwicklung von Pharmaca, zum Beispiel zur Förderung oder Hemmung der katalytischen Wirkung der codierten Proteine der Formeln I oder II.

5

25. Verwendung der Proteine mit den codierten Aminosäuresequenzen der Verbindungen der Formeln I oder II, einschliesslich der Proteine mit den separaten Teilsequenzen der codierten Aminosäuresequenzen der Verbindungen der Formeln I oder II, wie beispielsweise die separaten katalytischen Domänen, einschliesslich der  
10 Proteine mit den codierten Sequenzen oder Teilsequenzen der den Verbindungen der Formeln I oder II entsprechenden Splice-Varianten, einschliesslich der Proteine mit den codierten Aminosäuresequenzen, oder Teilsequenzen davon, der den Verbindungen der Formeln I oder II entsprechenden Allele, einschliesslich aller Sequenzvarianten der codierten Sequenzen, oder Teilen davon, der Verbindungen  
15 der Formeln I oder II, deren biologische Aktivität derjenigen der codierten Sequenzen der Verbindungen der Formeln I oder II gleich oder ähnlich ist, beispielsweise Sequenzvarianten der Verbindungen der Formeln I oder II, welche an den nicht konservierten Aminosäuresequenzpositionen von der Sequenz der Formeln I oder II abweichen, einschliesslich der Translationsprodukte der an die codierenden  
20 Sequenzen der Verbindungen der Formeln I oder II, oder an Teile davon, unter stringenten Bedingungen hybridisierenden Sequenzen, einschliesslich der Spezies-homologen Proteine der Verbindungen der Formeln I oder II, wie beispielsweise die Spezies-homologen Proteine der Ratte, des Kaninchens, des Rindes, des Schafes, des Schweins, der Primaten, der Vögel, des Zebrafisches, der Taufliege (*Drosophila melanogaster*), etc., einschliesslich der Teilsequenzen davon, wie beispielsweise die  
25 separaten katalytischen Domänen, für die Raumstrukturbestimmung, zum Beispiel der Raumstrukturbestimmung mittels Kristallographie oder Kernresonanzspektroskopie.

30

26. Verwendung der codierten Aminosäuresequenzen der Verbindungen der Formeln I oder II, einschliesslich der separaten Teilsequenzen der codierten Aminosäuresequenzen der Verbindungen der Formeln I oder II, wie beispielsweise die codierten Aminosäuresequenzen der separaten katalytischen Domänen der

Verbindungen der Formeln I oder II, einschliesslich der codierten Sequenzen oder Teilsequenzen der den Verbindungen der Formeln I oder II entsprechenden Splice-Varianten, einschliesslich der codierten Aminosäuresequenzen, oder Teilen davon, der den Verbindungen der Formeln I oder II entsprechenden Allele, einschliesslich  
5 aller Sequenzvarianten der codierten Sequenzen, oder von Teilen davon, der Verbindungen der Formeln I oder II, deren biologische Aktivität derjenigen der codierten Sequenzen der Verbindungen der Formeln I oder II gleich oder ähnlich ist, beispielsweise Sequenzvarianten der Verbindungen der Formeln I oder II, welche an den nicht konservierten Aminosäuresequenzpositionen von der Sequenz der  
10 Formeln I oder II abweichen, einschliesslich der Aminosäuresequenzen der Translationsprodukte der an die codierenden Sequenzen der Verbindungen der Formeln I oder II, oder an Teile davon, unter stringenten Bedingungen hybridisierenden Sequenzen, einschliesslich der Sequenzen der Spezies-homologen Verbindungen der Verbindungen der Formeln I oder II, wie beispielsweise die  
15 Sequenzen der Spezies-homologen Verbindungen der Ratte, des Kaninchens, des Rindes, des Schafes, des Schweins, der Primaten, der Vögel, des Zebrafisches, der Taufliege (*Drosophila melanogaster*), etc.) einschliesslich der Teilsequenzen der Spezies-homologen Verbindungen, wie beispielsweise die Sequenzen der katalytischen Domäne der Spezies-homologen Verbindungen, für Vorhersagen der  
20 Proteinstruktur mittels computerisierter Proteinstruktur-Vorhersage-Verfahren.

27. Verwendung der Raumstruktur der codierten Aminosäuresequenzen der Verbindungen der Formeln I oder II, einschliesslich der Raumstrukturen der separaten Teilsequenzen der Verbindungen der Formeln I oder II, wie beispielsweise  
25 die Raumstruktur der katalytischen Domäne, einschliesslich der Raumstruktur der codierten Sequenzen oder Teilsequenzen der den Verbindungen der Formeln I oder II entsprechenden Splice-Varianten, einschliesslich der Raumstruktur der codierten Sequenzen, oder Teilsequenzen, der den Verbindungen der Formeln I oder II entsprechenden Allele, einschliesslich der Raumstruktur aller Sequenzvarianten der  
30 codierten Sequenzen, oder von Teilen davon, der Verbindungen der Formeln I oder II, deren biologische Aktivität derjenigen der codierten Sequenzen der Verbindungen der Formeln I oder II gleich oder ähnlich ist, beispielsweise Sequenzvarianten der Verbindungen der Formeln I oder II, welche an den nicht konservierten

Aminosäuresequenzpositionen von der Sequenz der Formeln I oder II abweichen, einschliesslich der Raumstrukturen der Translationsprodukte der an die codierenden Sequenzen der Verbindungen der Formeln I oder II, oder an Teile davon, unter stringenten Bedingungen hybridisierenden Sequenzen, einschliesslich der  
5 Raumstrukturen der Spezies-homologen Verbindungen der Verbindungen der Formeln I oder II, wie beispielsweise die Raumstruktur der Spezies-homologen Verbindungen, oder von Teilen davon, der Ratte, des Kaninchens, des Rindes, des Schafes, des Schweins, der Primaten, der Vögel, des Zebrafisches, der Taufliege (*Drosophila melanogaster*), etc., als Zielstruktur für die Entwicklung von Pharmaca,  
10 zum Beispiel zur Hemmung oder Förderung der katalytischen Wirkung der codierten Proteine der Formeln I oder II.

28. Verwendung der codierenden Nucleotidsequenzen der Verbindungen der Formeln I oder II, einschliesslich der separaten Teilsequenzen der codierenden  
15 Sequenzen dieser Verbindungen der Formeln I oder II, wie zum Beispiel die codierenden Sequenzen der katalytischen Domänen der Verbindungen der Formeln I oder II, einschliesslich der codierenden Sequenzen oder Teilsequenzen der den Verbindungen der Formeln I oder II entsprechenden Splice-Varianten, einschliesslich der codierenden Sequenzen, oder Teilsequenzen, der den Verbindungen der Formeln  
20 I oder II entsprechenden Allele, einschliesslich aller Sequenzvarianten der codierenden Sequenzen, oder von Teilen davon, der Verbindungen der Formeln I oder II, deren Translationsprodukte in ihrer biologischen Aktivität denjenigen der Verbindungen der Formeln I oder II gleich oder ähnlich sind, beispielsweise Sequenzvarianten der Verbindungen der Formeln I oder II, welche an den nicht  
25 konservierten Aminosäuresequenzpositionen von der Sequenz der Verbindungen der Formeln I oder II abweichen, einschliesslich der an die codierenden Sequenzen, oder an Teile davon, unter stringenten Bedingungen hybridisierenden Sequenzen, einschliesslich der Nucleotidsequenzen, welche die durch die Verbindungen der Formeln I oder II codierten Proteine, oder Teile davon, codieren, aber, als Resultat  
30 unterschiedlicher Verwendung alternativer Codons, degeneriert sind bezüglich der als Verbindungen der Formel I oder II definierten Nucleotidsequenzen, in gentherapeutischen Anwendungen bei Menschen und bei Tieren, wie beispielsweise als Teile von Gentherapie-Vektoren oder wie beispielsweise als Teile von künstlichen

Chromosomen.

29. Verwendung der codierenden Nucleotidsequenzen der Verbindungen der Formeln I oder II, einschliesslich der separaten Teilsequenzen der codierenden Sequenzen dieser Verbindungen der Formeln I oder II, wie zum Beispiel die codierenden Sequenzen der katalytischen Domänen der Verbindungen der Formeln I oder II, einschliesslich der codierenden Sequenzen oder Teilsequenzen der den Verbindungen der Formeln I oder II entsprechenden Splice-Varianten, einschliesslich der codierenden Sequenzen oder Teilsequenzen der den Verbindungen der Formeln I oder II entsprechenden Allele, einschliesslich aller Sequenzvarianten der codierenden Sequenzen, oder von Teilen davon, der Verbindungen der Formeln I oder II, deren Translationsprodukte in ihrer biologischen Aktivität denjenigen der Verbindungen der Formeln I oder II gleich oder ähnlich sind, beispielsweise Sequenzvarianten der Verbindungen der Formeln I oder II, welche an den nicht konservierten Aminosäuresequenzpositionen von der Sequenz der Formeln I oder II abweichen, einschliesslich der an die codierenden Sequenzen, oder Teile davon, unter stringenten Bedingungen hybridisierenden Sequenzen, einschliesslich der Nucleotidsequenzen, welche die durch die Verbindungen der Formeln I oder II codierten Proteine, oder Teile davon, codieren, aber, als Resultat unterschiedlicher Verwendung alternativer Codons, degeneriert sind bezüglich der als Verbindungen der Formel I oder II definierten Nucleotidsequenzen, für sogenannte Cell-Engineering-Anwendungen zur Produktion von gentechnologisch veränderten Zellen, welche die codierten Sequenzen, oder Teile davon, der Verbindungen der Formeln I oder II produzieren, zum Beispiel zum Zweck zelltherapeutischer Anwendung als Medikament nach Anspruch 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, oder 21.

30. Verwendung der codierten Aminosäuresequenzen der Verbindungen der Formeln I oder II, einschliesslich der separaten Teilsequenzen der codierten Aminosäuresequenzen der Verbindungen der Formeln I oder II, wie beispielsweise die codierte Aminosäuresequenz der katalytischen Domäne oder einer oder mehrerer der andern Domänen oder Segmente, einschliesslich der codierten Sequenzen oder Teilsequenzen der den Verbindungen der Formeln I oder II entsprechenden Splice-Varianten, einschliesslich der codierten Sequenzen oder Teilsequenzen der den

Verbindungen der Formeln I oder II entsprechenden Allele, einschliesslich aller Sequenzvarianten der codierten Sequenzen, oder Teilen davon, der Verbindungen der Formeln I oder II, deren biologische Aktivität derjenigen der codierten Sequenzen der Verbindungen der Formeln I oder II gleich oder ähnlich ist, beispielsweise

5 Sequenzvarianten der Verbindungen der Formeln I oder II, welche an den nicht konservierten Aminosäuresequenzpositionen von der Sequenz der Formeln I oder II abweichen, einschliesslich der Translationsprodukte, oder Teilen davon, der an die codierenden Sequenzen der Verbindungen der Formeln I oder II, oder an Teile davon, unter stringenten Bedingungen hybridisierenden Sequenzen, einschliesslich

10 der codierten Sequenzen der Spezies-homologen Verbindungen der Verbindungen der Formeln I oder II, wie beispielsweise die codierten Sequenzen der Spezies-homologen Verbindungen der Ratte, des Kaninchens, des Rindes, des Schafes, des Schweins, der Primaten, der Vögel, des Zebrafisches, der Taufliege (*Drosophila melanogaster*), etc., einschliesslich der separaten Teilsequenzen der codierten

15 Sequenzen der Spezies-homologen Verbindungen der Verbindungen der Formeln I oder II, wie beispielsweise die codierte Aminosäuresequenz der katalytischen Domäne, oder einer oder mehrerer der andern Domänen oder Segmente, als Antigene zur Herstellung von Antikörpern, wie beispielsweise Antikörper, welche die Protease-Funktion hemmen oder fördern, oder Antikörper, welche für

20 immunohistochemische Studien eingesetzt werden können.

31. Verwendung der codierenden Nucleotidsequenzen der Verbindungen der Formeln I oder II, einschliesslich der separaten Teilsequenzen der codierenden Sequenzen dieser Verbindungen der Formeln I oder II, wie zum Beispiel die

25 codierenden Sequenzen der katalytischen Domänen der Verbindungen der Formeln I oder II, einschliesslich der codierenden Sequenzen oder Teilsequenzen der den Verbindungen der Formeln I oder II entsprechenden Splice-Varianten, einschliesslich der codierenden Sequenzen, oder Teilsequenzen, der den Verbindungen der Formeln I oder II entsprechenden Allele, einschliesslich aller Sequenzvarianten der

30 codierenden Sequenzen, oder von Teilen davon, der Verbindungen der Formeln I oder II, deren Translationsprodukte in ihrer biologischen Aktivität denjenigen der Verbindungen der Formeln I oder II gleich oder ähnlich sind, beispielsweise Sequenzvarianten der Verbindungen der Formeln I oder II, welche an den nicht



- 5 konservierten Aminosäuresequenzpositionen von der Sequenz der Formeln I oder II abweichen, einschliesslich der an die codierenden Sequenzen, oder an Teile davon, unter stringenten Bedingungen hybridisierenden Sequenzen, einschliesslich der Nucleotidsequenzen, welche die durch die Verbindungen der Formeln I oder II
- 10 codierten Proteine, oder Teile davon, codieren, aber, als Resultat unterschiedlicher Verwendung alternativer Codons, degeneriert sind bezüglich der als Verbindungen der Formel I oder II definierten Nucleotidsequenzen, für die Herstellung von transgenen Tieren, wie beispielsweise transgene Mäuse.
- 15 32. Verwendung der codierenden Nucleotidsequenzen der Verbindungen der Formeln I oder II, einschliesslich der separaten Teilsequenzen der codierenden Sequenzen dieser Verbindungen der Formeln I oder II, wie zum Beispiel die codierenden Sequenzen der katalytischen Domänen der Verbindungen der Formeln I oder II, einschliesslich der codierenden Sequenzen oder Teilsequenzen der den
- 20 Verbindungen der Formeln I oder II entsprechenden Splice-Varianten, einschliesslich der codierenden Sequenzen, oder Teilsequenzen, der den Verbindungen der Formeln I oder II entsprechenden Allele, einschliesslich aller Sequenzvarianten der codierenden Sequenzen, oder von Teilen davon, der Verbindungen der Formeln I oder II, deren Translationsprodukte in ihrer biologischen Aktivität denjenigen der
- 25 Verbindungen der Formeln I oder II gleich oder ähnlich sind, beispielsweise Sequenzvarianten der Verbindungen der Formeln I oder II, welche an den nicht konservierten Aminosäuresequenzpositionen von der Sequenz der Formeln I oder II abweichen, einschliesslich der an die codierenden Sequenzen, oder Teile davon, unter stringenten Bedingungen hybridisierenden Sequenzen, einschliesslich der
- 30 Nucleotidsequenzen, welche die durch die Verbindungen der Formeln I oder II codierten Proteine, oder Teile davon, codieren, aber, als Resultat unterschiedlicher Verwendung alternativer Codons, degeneriert sind bezüglich der als Verbindungen der Formel I oder II definierten Nucleotidsequenzen, für die Inaktivierung oder Abänderung des entsprechenden Gens mittels gezielter Eingriffe am Gen durch sogenannte "Gene Targeting"-Techniken, wie beispielsweise der Elimination des Gens bei der Maus durch homologe Rekombination.

33. Verwendung der codierenden Nucleotidsequenzen der Verbindungen der

Formeln I oder II, einschliesslich der separaten Teilsequenzen der codierenden Sequenzen dieser Verbindungen der Formeln I oder II, wie zum Beispiel die codierenden Sequenzen der katalytischen Domänen der Verbindungen der Formeln I oder II, einschliesslich der codierenden Sequenzen oder Teilsequenzen der den Verbindungen der Formeln I oder II entsprechenden Splice-Varianten, einschliesslich der codierenden Sequenzen, oder Teilsequenzen, der den Verbindungen der Formeln I oder II entsprechenden Allele, einschliesslich aller Sequenzvarianten der codierenden Sequenzen, oder von Teilen davon, der Verbindungen der Formeln I oder II, deren Translationsprodukte in ihrer biologischen Aktivität denjenigen der Verbindungen der Formeln I oder II gleich oder ähnlich sind, beispielsweise Sequenzvarianten der Verbindungen der Formeln I oder II, welche an den nicht konservierten Aminosäuresequenzpositionen von der Sequenz der Formeln I oder II abweichen, einschliesslich der an die codierenden Sequenzen, oder Teile davon, unter stringenten Bedingungen hybridisierenden Sequenzen, einschliesslich der Nucleotidsequenzen, welche die durch die Verbindungen der Formeln I oder II codierten Proteine, oder Teile davon, codieren, aber, als Resultat unterschiedlicher Verwendung alternativer Codons, degeneriert sind bezüglich der als Verbindungen der Formeln I oder II definierten Nucleotidsequenzen, zur Diagnostik von Störungen im Gen, welches der Verbindung der Formel I zu Grunde liegt.

20

34. Verwendung der codierenden Nucleotidsequenzen der Verbindungen der Formeln I oder II, einschliesslich der separaten Teilsequenzen der codierenden Sequenzen dieser Verbindungen der Formeln I oder II, wie zum Beispiel die codierenden Sequenzen der katalytischen Domänen der Verbindungen der Formeln I oder II, einschliesslich der codierenden Sequenzen oder Teilsequenzen der den Verbindungen der Formeln I oder II entsprechenden Splice-Varianten, einschliesslich der codierenden Sequenzen oder Teilsequenzen der den Verbindungen der Formeln I oder II entsprechenden Allele, einschliesslich aller Sequenzvarianten der codierenden Sequenzen, oder von Teilen davon, der Verbindungen der Formeln I oder II, deren Translationsprodukte in ihrer biologischen Aktivität denjenigen der Verbindungen der Formeln I oder II gleich oder ähnlich sind, beispielsweise Sequenzvarianten der Verbindungen der Formeln I oder II, welche an den nicht konservierten Aminosäuresequenzpositionen von der Sequenz der Formeln I oder II abweichen,

einschliesslich der an die codierenden Sequenzen, oder an Teile davon, unter stringenten Bedingungen hybridisierenden Sequenzen, einschliesslich der Nucleotidsequenzen, welche die durch die Verbindungen der Formeln I oder II codierten Proteine, oder Teile davon, codieren, aber, als Resultat unterschiedlicher

5 Verwendung alternativer Codons, degeneriert sind bezüglich der als Verbindungen der Formel I oder II definierten Nucleotidsequenzen, als Ausgangssequenz für gentechnische Modifikationen zum Zweck der Erzeugung von Medikamenten oder Gentherapie-Vektoren, welche im Vergleich zu entsprechenden Medikamenten oder Gentherapie-Vektoren, welche die codierende Nucleotidsequenz der Verbindungen

10 der Formel I oder II enthalten, veränderte Eigenschaften haben, wie beispielsweise veränderte proteolytische Aktivität, veränderte proteolytische Spezifität, oder veränderte pharmakokinetische Eigenschaften.

### Zusammenfassung

Es werden Neurotrypsine der Formeln I oder II, einschliesslich die separaten  
codierenden und codierten Sequenzen dieser Verbindungen der Formeln I oder II  
5 beschrieben.

Diese Verbindungen können als wenigstens eine aktive Verbindung in einem  
Medikament verwendet werden.

10 Die codierten Peptidsequenzen dieser Verbindungen können als  
Zielsubstanzen (Targets) für die Entwicklung von Pharmaca verwendet werden.